

From THE CENTRE FOR VIOLENCE PREVENTION,  
DEPARTMENT OF CLINICAL NEUROSCIENCE  
Karolinska Institutet, Stockholm, Sweden

# **STRUCTURED CLINICAL ASSESSMENT AND MANAGEMENT OF RISK OF VIOLENT RECIDIVISM IN MENTALLY DISORDERED OFFENDERS**

*Mats Dernevik*



Stockholm 2004



All previously published papers were reproduced with permission from the publisher.

Published and printed by Karolinska University Press  
Box 200, SE-171 77 Stockholm, Sweden  
© Mats Dernevik, 2004  
ISBN 91-7140-095-8



# ABSTRACT

*Dernevik, M. (2004). Structured clinical assessment and management of risk of violent recidivism in mentally disordered offenders. Stockholm, Sweden: Karolinska Institutet, department of Clinical Neuroscience, Centre for Violence Prevention. ISBN 91-7140-095-8*

**Background:** The assessment of risk of violence among mentally disordered offenders has been a controversial and well-researched area in forensic psychology and psychiatry in the last decades. The main focus of this research has been on the predictive validity of various risk factors and methods of combining risk factors to gain the highest possible predictive accuracy. In the present thesis, risk assessment is defined more broadly than predictive accuracy, and also includes process factors, risk-management and the communication and decision-making associated with risk of violence. The overall aim of this thesis was to study was to explore the process of structured risk assessment in its naturalistic clinical setting.

**Method:** Four different samples were included in the 5 studies. A guideline for structured clinical risk assessment, Historical-Clinical-Risk assessment (HCR-20), was used in the first 4 studies. The first study used a 6 raters x 6 patients design to establish inter-rater reliability and validity of the HCR-20. 54 forensic patients were followed over time and monitored for inpatient violence and violence after discharge during three risk-management conditions in study 2. A sample of 40 nurses, assessing the same 8 patients, was included in studies 3 and 4. Finally study 5 included a sample of 88 decision-makers, divided into 3 groups; Clinicians, Criminal law professionals and Controls.

**Results:** The HCR-20 was found to have reasonable reliability and validity in study 1. The main finding in study 2 was that the predictive accuracy of the HCR-20 was influenced by the intensity of risk management (AUC .64 compared to .82). In study 3 we found that structured clinical risk assessment was not “immune” to emotional bias in the assessment process. 43% of the variance in risk-scores could be attributed to the assessors’ emotions towards the patient. The information utilised to make the assessment and how the assessor values it, also influenced the assessments in study 4. Placing value on personal interaction was more associated with inpatient violence than with recidivism. In study 5 we found that the inclination of making release decisions was greatly influenced ( $\eta = .58$ ) by the prospect of making false negative error of judgement.

**Conclusions:** Structured clinical risk assessments can be undertaken in a reliable and valid way in forensic clinical settings. Attention needs to be paid to factors that might influence the outcome of the assessments and the risk-management decisions that are the consequence of risk assessment. These factors can be emotional biases, evaluation of different kinds of information that form the basis for the process. There also needs to be an awareness of other factors than probabilities that influence decisions about risk. It is suggested that future descriptive, as opposed to prescriptive, research is needed on the processes and influences on risk assessments as they are actually done by clinicians in forensic, psychiatric and correctional settings.



## LIST OF PUBLICATIONS

- I. Dernevik, M. (1998). Preliminary findings on reliability and validity of the Historical- Clinical-Risk assessment in a forensic psychiatric setting, *Psychology, Crime & Law*, 4, 127-137.
- II. 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 & Law*, 8, 93-111.
- III. Dernevik, M., Falkheim, M., Holmqvist, R., & Sandell, R. (2001). Implementing Risk Assessment Procedures in a Forensic Psychiatric Setting: Clinical Judgement Revisited. In: D. Farrington, C. Hollin, & M. McMurrin, *Sex and Violence: The psychology of crime and risk assessment*. London, Routledge, pp. 83-101.
- IV. Dernevik, M. (2004). *Professional risk assessment in clinical practice: The role of relevant clinical data gathering and context for the accuracy of structured professional risk assessment*. Manuscript. Stockholm: Centre for Violence prevention, Karolinska Institutet.
- V. Dernevik, M., & Grann, M. (2004). *Factors influencing release decision-making: An experimental study*. Manuscript. Stockholm: Centre for Violence prevention, Karolinska Institutet.

# CONTENTS

Abstract .....	5
List of Publications .....	7
List of Abbreviations .....	9
<b>BACKGROUND</b> .....	11
Violence and mental disorder .....	13
Violence Prediction and Risk Assessment .....	15
Defining Risk Assessment .....	20
Ethical considerations in risk prediction .....	22
Methods of Risk Assessment .....	23
Implementation into Clinical Practice .....	29
Risk communication .....	31
Risk Decision-making .....	32
<b>AIMS</b> .....	35
<b>METHODS</b> .....	37
Designs and Procedures .....	40
Statistics .....	43
Research Ethical considerations .....	44
<b>RESULTS</b> .....	47
<b>DISCUSSION</b> .....	51
Limitations .....	51
Comments to Results .....	52
Discussion Proper .....	54
Conclusions .....	56
<b>ACKNOWLEDGEMENTS</b> .....	61
<b>REFERENCES</b> .....	63

## LIST OF ABBREVIATIONS

ATMDO	Attitudes to Mentally Disordered Offenders
AUC	Area Under the Curve
BRÅ	Brottförebyggande Rådet
DSM	Diagnostic and Statistical Manual of Mental Disorders
FPE	Forensic Psychiatric Evaluation
FWL	Feeling Word list
HCR-20	Historical-Clinical- Risk assessment scheme
ICD	International Classification
Kval	Kriminalvårdslagen
NPV	Negative Predictive Value
PCL-R	Psychopathy Checklist-Revised
PCL:SV	Psychopathy Checklist: Screening Version
PPV	Positive Predictive Value
ROC	Receiver Operating Characteristics
SOAS	Staff Observation Aggression Scale
SORM	Structured Outcome Assessment & Community Risk Monitoring
TCO	Threat Control Override
VRAG	Violence Risk Appraisal Guide



## **BACKGROUND**

**T**he fear of crime is generally increasing in society and reports in media and the public debate frequently concern violent crime. Leading up to elections, "law and order" is typically one of the top five topics of public concern and interest. In a public survey leading up to the Swedish election in 2002, crime in general was rated in fourth place of general concerns, just after schools, child-care and health-care (Friberg, 2003)

Parallel to the public debate there has been scientific investigations into the area from behavioural and political scientists. A major topic is whether violent crime is truly increasing or whether the tolerance for violence is decreasing in the general public. A decreased tolerance would result in people being more inclined to report violent behaviour to the police and hence explain the growth of reported violent crime, evident in crime-statistics in Sweden (BRÅ, 2002).

Violent crime could be considered in parallel to sexual crime. Few people believe that sexual abuse of children or rapes of women actually have increased in recent decades, despite a sharp increase in the number of reported crime. The perception of criminologists and political scientists is that, the inclination to report these types of crimes have increased, while the true frequency of sexual crime remains unknown. The difference between the actual frequency of crime and the number of reported crimes is the hidden number of crime. There is some evidence that the hidden number for rape has decreased in recent years (e.g. Blackburn, 1993). However, the number of reported rapes in Sweden, more than doubled between 1975 and 2001, when approximately 2100 cases, nationally, were reported to the police (BRÅ, 2002).

A similar reasoning to that of hidden numbers in sexual crime could be applied to the area of violent crime. The example of rape is indeed both a sexual and violent crime. Non-sexual, violent assaults constitute a majority of violent crimes reported to the Swedish police. The number of reported violent assaults has almost tripled in 26 years, from approx. 24.000 in 1975 to almost 70.000 cases in 2001 (BRÅ, 2002).

Four particularly frightening and un-provoked violent crimes during the summer of 2003 caused the fear of unprovoked violence from mentally disordered offenders to soar in Sweden. Among these crimes was the murder of the Swedish Foreign Secretary. There seems to be a widespread perception among the public that this kind of violence has become more frequent and more serious in character. There are, however, at least two findings that would suggest that this might not be entirely true. The Swedish council for Crime Prevention suggested that; "physical assaults, where perpetrator and victim know each other is one of the types of crime that may have increased the inclination to report crimes" (BRÅ, 2002, p.23). If the tolerance for domestic and other types of violence, where there is some kind of relationship between assailant and victim, might have decreased this may account for the rise in reported violent crime. The second point that rebukes the increase in violence theory is the lack of increase in the most serious form of violent crime; murder and manslaughter. The

Swedish statistics (BRÅ, 2002) show no, or a very slight, increase murder and manslaughter convictions from 1975 to 2001. According to another source; the Swedish causes of deaths register, kept by the department of medicine and welfare (Socialstyrelsen), and unrelated to police reports, there is no increase in violent deaths at all. If street violence truly were on the increase and indeed getting more brutal and serious, the argument would be that this would show as an increase in this register (Von Hofer, 2000).

Whether actually increasing or not, violence is one of the core problems in society. In recent years it has increasingly been considered as a public health issue (Folhälsoinstitutet, 1995; Melinder, 2002). Criminal violence is a major cause for loss of life worldwide (Murray & Lopez, 1997). In fact, the World Health Organisation recently reported that criminal violence causes twice as many casualties compared to armed conflicts worldwide (1424 versus 849 deaths per average 24 hours) (WHO, 2002). This is only considering deaths. The suffering of victims of violence who survive, and witnesses, victims' family members etc. has not been estimated. Apart from these effects there is the growing fear of criminal violence (Bilsky & Winkel, 2002; Gabriel, 2002) and the loss of freedom, sense of security and faith in the general public. All these factors make the economical impact of violent crime very high, but difficult to estimate. Among the costs for society are of course the cost of medical and psychological care and the loss of income from work for the victims of violence. The cost of police investigations, prosecution, lawyers and courts dealing with violent offenders should be considered as well as the cost of the correctional system, whether in the form of prison services, forensic mental health and other forms of punitive and rehabilitation services. On a secondary level, the fear of violent crime causes measures to be taken, that are unproductive to the quality of life of citizens in the form of security systems, surveillance of public and private facilities, alarms etc.

From a scientific view violence is a very complex problem and the subject of a vast array of study. Violence has been studied from many aspects, from biomedical levels, considering the role of neuron transmitters and hormonal balance, to the macro-level of sociology and political science.

Psychology as a discipline, perhaps placed between the basic and macro-levels, has made major contributions to the study of the mechanisms of violence. Starting 50 years ago with the classical studies on aggression by Milgram (1963), Berkowitz (1964), and Bandura (1973), social learning theory has been the main framework of understanding of human aggression and propensity for violence.

Despite the psychological study of man's general propensity for violence, studies suggest that it is a relatively limited number of individuals who are perpetrating a disproportional part of criminal, interpersonal violence (Farrington, 1996; Stattin & Magnusson, 1991). There is some evidence as to what kind of factors are associated to this group of predominantly men, but the knowledge base is incomplete and the field is still developing. Many of the characteristics of these men are described on an individual level and the psychological focus is on individual differences. These psychological factors include attitudes to violence and other cognitive schemas, affective processing, temperamental and personality features such as impulsivity, egocentricity and poor

problem-solving skills (McGuire, 2000). Research has also shown that there is little or no support for risk factors previously believed to be strongly associated with violent crime, such as low socio-economic status (SES), low self-esteem, or mental disorder (See section Violence and mental disorder, below).

Researchers and theorists have argued that a risk factor approach, in the same way as the concept of risk is typically utilised in epidemiological health research, is a useful approach also to violent crime (Andrews & Bonta, 1998). Andrews (1995, p. 36) provided a rationale:

” Individuals varying in their criminal past (as documented by cross-sectional studies) and their criminal future (as documented in longitudinal studies) may be differentiated at levels well above chance on a number of situational, circumstantial personal, interpersonal, familial and structural / cultural/economic factors.”

With this approach it is possible to empirically address the question: What individuals have a high risk of violent re-offending. Or, more specifically: Which are the individuals, out of an identified population of violent criminals, who share characteristics with reasonably similar members of other well-studied groups that have *de facto* recidivated. Further research could establish which of these factors are merely markers and which are true, causal risk factors (Kraemer et al., 1997). This, in turn, would also make it possible to identify the factors that actually cause the increased risk in these individuals. To the degree that these factors are dynamic or amenable to change, or at least manageable, strategies of risk management and prevention of violent crime would also be possible.

## **VIOLENCE AND MENTAL DISORDER**

This thesis is based on a population of individuals who have been convicted and sentenced to forensic psychiatric care according to the Swedish Law on Forensic Psychiatric Care (FPC) of 1992. The relationship between violence and mental disorder is complex and not without controversy. Epidemiological research has in the past shown that major mental disorder is associated with an elevated risk of violent crime (Brennan et al., 2000; Eronen et al., 1996; Hodgins, et al., 1996; Lindqvist & Allebeck, 1990; Steadman et al., 1998; Swanson, 1994; Tiihonen et al., 1997). However, looking at this from another perspective, only a minor part of all violent crimes are committed by persons with a mental disorder. In a recent study, Fazel and Grann (2004a) found an association between violent crime and having been admitted to hospital with a psychiatric diagnosis. They studied all 324,000 violent crimes during 1988-2000 in Sweden, and identified all 441,000 persons who during at least ones this period were hospitalised with a psychiatric diagnosis. Persons who, during the study period, had received psychiatric care committed about a third of violent crimes. For more serious violent crime, murder and manslaughter, the proportion of mental disordered offenders was even higher. However, the contribution of substance abuse disorders was high, while the classical major mental disorders had a low representation in violent crime. For example were people with a diagnosis of schizophrenia involved in only 2.5% of violent crimes (see also: Grann & Fazel, 2004; and Fazel & Grann, 2004b). Findings like these notwithstanding; there is a

prevailing fear for this type of perpetrator with a major mental disorder. In media and entertainment, the stereotype of the enigmatic psychotic killer is depicted as guilty of heinous and unpredictable violence to innocent and unknown victims. In the public perception, the risk of violence in mentally disordered offenders is over-estimated (Redendo et al., 1996). There is also evidence that the fear of violence from mentally disordered offenders is increasing in Europe and North America (Main et al., 1996). Contrary to these fears and the stereotypes of psychotic killers in Hollywood, the reality is often considerably less dramatic. This type of offender attacking unknown victims is in fact very rare (Nordström & Kullgren, 2003; O’Kane & Bentall, 2000).

Many studies points to an overlap between criminal and psychiatric populations. In an early and typical such study Gunn and colleagues (1978) found that more than 20% of a large prison population were suffering from clear and serious mental disorders. The most frequent diagnoses were mood and anxiety disorders. This prevalence obviously exceeds that of the general population. In a sample of a sample of UK prisoners with lifetime sentences, 10% suffered from schizophrenia (Taylor, 1986). A recent meta-analysis of 62 studies from 12 countries and including 23000 prisoners, on mental health in prison, suggested that the prevalence of antisocial personality disorder was 46%, and the prevalence of major depression was 12%, and 4% suffered from psychotic illnesses (Fazel & Danesh, 2002).

A higher-than-expected prevalence of mental disorder in prison populations does not, however, imply a causal relationship between mental disorder and criminality. Long imprisonment itself might lead to psychiatric complaints or persons with mental disorder might have a higher risk of detection, have a higher propensity to admit or confess to crimes or be subject to longer sentences than criminals with no psychiatric disorders, whose criminal behaviour might be less likely to be detected, prosecuted and sentenced (Gunn et al., 1978). The association between, for instance, schizophrenia and violent crime should not be interpreted as a causal link between disorder and crime. Although studies have found that the disorder typically comes before the criminal behaviour in time (Taylor & Hodgins, 1994), a considerable number of mentally disordered offenders have committed crimes before receiving a diagnoses and entering into the psychiatric services. Tengström and Hodgins (2002), while maintaining the belief that psychotic disorder and criminal offending are causally linked, called these patients “early starters”. Munkner and colleagues (2003) found that 27% of psychiatric patients had committed a crime before the first psychiatric hospital contact in Denmark. There is some evidence to assume that there are common, underlying factors both in the clinical presentation of mental disorder and in criminal behaviour. These clinical factors and criminogenic factors might have a considerable overlap (Blackburn, 1993). On this common ground could be socio-economic factors (Farrington, 1992), poor child rearing, and school performance (Farrington 1995). Regardless of aetiology there seems to be higher than expected frequencies of criminal behaviours, in psychiatric populations (Howells, 1982; Prins, 1986).

From the forensic perspective, mental disorder has sometimes been considered a non-correlate of violent crime (Monahan, 1984). More recently Monahan and colleagues (2001) found the base rate of violent behaviour in psychiatric outpatients to be high. This USA study found 4 murders among 970 civil psychiatric patients during a

6 months study period. In the same study; *The MacArthur Violence Risk Assessment Study*, also found a diagnosis of schizophrenia to be negatively associated to the risk of violent behaviour. In accordance to this finding, Quinsey and colleagues previously (1998) found that a diagnosis of schizophrenia was a protective factor for the risk of violent recidivism in a Canadian forensic psychiatric sample. These findings might seem counterintuitive and in opposition to other studies such as Lindqvist and Allebeck (1990) and Hodgins and colleagues (1996). However it is important to consider the nature of the attrition in the samples used. Comparing schizophrenics to other groups of mentally disordered offenders, who are probably predominantly personality-disordered, which in turn was found to be strongly risk of violence in both studies above.

Major mental disorder and symptoms are considered to be risk factors of violence in most risk assessment schemes. In the HCR-20, the authors (Webster et al., 1997; Douglas et al., 2001) describe mental health items both in the historical and clinical factors sections (for a description of the HCR-20 scheme, see section Structured clinical methods, below). Having been diagnosed with major mental disorder (DSM IV, axis I) and personality disorder (axis II), ever in life are considered historical factors, increasing the individuals' risk of future violence. In the clinical items section current symptoms of major mental disorder is viewed as enhancing risk. There have been some suggestions about the mechanisms about active symptoms and risk of violence. The affective arousal associated with hallucinations has been shown to be predictive of risk of imminent violence (Nijman & Palmstierna, 2002; Bowers et al., 2002). Certain type of hallucinations, Threat-Control-Override (TCO) has been the subject for examination for links with violence (Link & Stueve, 1994; Björkly & Havik, 2003). Paranoid delusions may, for obvious reasons, lead to violence (Baxter, 1997). With an innovative, qualitative design, Nordström and colleagues (2004) examined a set of 48 consecutive court-referred homicide cases of persons with schizophrenia. Nordström et al. interpreted their observations to suggest that around half of the homicides of the violent acts were, in fact, related to paranoid ideation.

## **VIOLENCE PREDICTION AND RISK ASSESSMENT**

### **History**

The controversy and shifting language surrounding the prediction of dangerousness is a fairly modern one. The question emerged in the late 1960's and early 1970's with two famous legal cases in the U.S.A. It seems that before this time it was generally assumed by judges and legislators that mental health and correctional professionals knew which individuals were prone to violence and who was not. Particularly the Baxtrom case from 1966 was instrumental in pointing the public debate and scientific attention to the assessment of dangerousness. In the wake of this case, two researchers, Steadman and Cocozza (1974) were able to report on almost 1,000 subjects, all deemed to be a danger to the public, that were released from the Danemora state hospital in the state of New York as a consequence of this "naturally occurring experiment" (Webster & Bailes, 2001). Johnny Baxtrom, a Swedish descendant, was institutionalised cause of his assumed dangerousness under legal

requisite of “psychopathic” disorder. After a political campaign for his case, the US high court in 1966 ruled that the institutionalisation was unlawful, since the state could not show that the assumption of expertise in those psychologists testifying that future violent behaviour could be substantiated. Thus, almost 1,000 “dangerous psychopaths” were turned loose on the streets of New England. The researchers were able to follow Baxtrom and some of his fellow ex-patients after release. They found that only 2 % were involved in further violence after 2-3 years (Webster & Bailes, 2001). In 1974, Steadman and Coccozza concluded in their book “Careers of the Criminally Insane” that clinicians tend to see dangerousness in too many of their patients. The study was, perhaps unsurprisingly, criticised both from agents in the legal system and by other researchers. The Baxtrom “release cohort” was older than most mentally disordered offenders (because they had been incarcerated for a long time) and it was argued that (high) age is a known protective factor for violent behaviour. However the results were more or less replicated following another case in the U.S., the “Dixon case”. Some years later, Thornberry and Jacoby (1979) were able to report similar outcome for a cohort of younger men in the state of Pennsylvania. At the time this became part of the starting signal for a political and scientific development that was highly critical of what was perceived as erratic and arbitrary decisions made by power-oriented experts. This development took place at a time when patients’ rights, interests and safety before the law were considered important issues and a zeitgeist that could be described as anti-psychiatric and anti-authoritarian (Rosenhan, 1973; Szasz, 1974)

An entire research tradition, on the reliability and validity of clinical judgement in psychiatry and clinical psychology, was established in the mid 1900’s by Paul Meehl and colleagues (1954; 1957). This research convincingly showed that clinicians couldn’t claim expertise as far as unstructured “gut feelings” judgements are concerned. As solid evidence amassed for the lack of validity of clinical judgements about dangerousness in mentally disordered offenders in the 1970s, Ennis and Litwack (1974) summarised the evidence in a paper, legendarily entitled “Psychiatry and the presumption of expertise: Flipping coins in the courtroom”. This article showed that clinicians who try to predict recidivism rarely do better than chance. In the 1970s, as is still the case today, mentally disordered individuals who have committed violent acts were perceived as generally dangerous. Decisions about the incarceration of these individuals are still made by courts, or a similar tribunal. In these legal contexts clinicians are still called upon for their expertise regarding the issue of dangerousness of these individuals. The clinicians, the vast majority of which are psychiatrists and clinical psychologists, still make as good and honest assessments as they can and communicate them to the Judge, or the tribunal, whose job it is to make decisions about denying liberties and freedom to mentally disordered offenders (Grann, 2002). The scientific development of the “first generation” of violence risk assessment research in the 1970s resulted in a multitude of studies showing the unreliability of unstructured clinical judgement that Ennis and Litwack so elegantly summarised in their 1974 paper. In the late 1970’s and the early 1980’s, many clinicians took the view that clinical assessments for dangerousness were unprofessional and unethical and that this kind of assessments should nor could be done in a responsible clinical practice, and that clinicians should abstain from making this kind of assessments of individuals (Stone, 1985; Appelbaum, 1997). Despite this professional pessimism, the task, in itself, of violence risk assessment could not be ignored. Shah (1978) argued that mental health professionals were obliged to submit

opinions on individuals' dangerousness in correctional, civil psychiatric and forensic psychiatric contexts. In an influential text, Monahan (1981) called for a "second generation" of research on the prediction of violent behaviour. The main argument was for actuarial assessments and for research that could demonstrate correlations between demographic, static, historical factors and future violent behaviour. In the 1980s, Monahan (1984; 1988) also called for prediction-outcome studies that with relatively brief follow-up periods, avoiding the recurrent problems with predictions of community violence, that were being made in institutional settings and with a large time-span and changing conditions (including clinical interventions, i.e. treatment) between prediction and outcome.

These recommendations came to act as a blueprint of a "second generation" of research and policy in risk assessment, which prevailed for two decades (Grann, 2002; Webster & Bailes, 2001).

During the 1990s some developments of clinical constructs have been introduced into the risk assessment literature. Probably the most prominent of these is the conceptualisation and operationalisation of psychopathy (Hare, 1991; 2003; Hart et al., 1995). The link between psychopathy and violent reoffending is robust and has been reported from a variety of samples and researchers (Grann et al., 1999; Hare et al., 2000; Hemphill et al., 1998; Monahan et al., 2000; Rice & Harris, 1995; Salekin et al., 1996; Skeem & Mulvey, 2001; Tengström et al., 2000; Walters, 2003). Among other clinical constructs that have been shown to be associated with violence risk are Impulsivity (Barrat, 1994; Webster & Jackson 1997) and Anger (Novaco, 1994).

## **Issues in Violence Prediction**

### *Clinical prediction of violence*

A major problem in the field of risk assessment is perhaps the lack of theories for what actually *causes* violent recidivism. The research described in this thesis is typically of a pragmatic-empiricist nature. There seems to be a scarcity of studies focusing on the aetiology of violent recidivism. Blackburn (2000) stated that:

"The lack of a universally accepted theoretical framework for understanding violence is the major limitation to theoretically-driven clinical assessments" (Blackburn, 2000; p. 196).

The possible exception may be the psychopathic personality disorder, for which the origins of the condition have to some extent been subject to scientific investigations (Paris, 1998; van Tijen & Verheul, 2002).

Despite the relative accuracy in prediction of violent recidivism demonstrated by the "second generation" research (Bonta et al., 1998; Borum, 1996; Dolan & Doyle, 2000), it is prudent to point out that the task of risk assessment has been limited. The task of being more accurate than chance, as set out by the very discouraging findings on unstructured clinical judgement (Ennis & Litwack, 1974; Meehl, 1954), is a fairly modest one. However modest, it seems to have been worthwhile judging from the large number of studies focusing on individual risk factors and striving to combine them in risk assessment procedures in the 1990s (Borum, 1996).

The emphasis in risk assessment has been on the use of statistically derived risk factors and on actuarial or highly structured ways of combining and interpreting the presence or absence of factors, all to avoid the subjectivity and arbitrary nature of clinical judgement. This is for the benefit of the legal rights of the client. However successful in demonstrating statistically significant predictive validity (better than chance) this approach has proven to be, it also carries some inherent limitations (Belfrage et al., 2000; Douglas et al., 1999; Grann & Nilstun, 2000). After all, there is a substantial difference between statistical and clinical significance (de Vet et al., 2003). Time is a problem in this context. A fundamental problem is the evaluation of accuracy of risk assessment without releasing individuals who are clearly deemed as dangerous into the community (Litwack, 2001). Most studies have approached this by using retrospective designs. Because release decisions cannot be randomised and controlled for research purposes and because commitment times in forensic psychiatry are usually long, the obvious way of obtaining data, is to look at samples of released individuals and then appraise how dangerous they were at the time of discharge. With some recent exceptions (Belfrage & Douglas, 2002; Dolan & Khawaja, 2004), prospective studies on risk assessment are few (Björkly, 1995). It is indeed paradoxical, and perhaps ironic, that a field dealing with prediction into the future is restricted to look backwards in time. The retrospective problem also means that the range of predictors, possible to examine, is limited.

The development has been characterised by “shotgun empiricism” (Blackburn, 2000) and “black box epidemiology” (Susser & Susser, 1996). The factors tend to be those that are readily available from records or which are subject to general criminological consensus (Gendreau et al., 1996). The factors also have a bias towards history. Researchers tend to favour factors that are easy to measure (albeit not necessarily heuristic) rather than those that are difficult to measure (but potentially valid, true causal factors) (Hart, 1998). This entails that factors tend to be seen as static traits rather than dynamic or even acute (Hanson & Bussière, 1998). The, sometimes fleeting, nature of aggression is typically not captured by static factors. By definition, this approach to risk assessment also minimises the contribution of the clinician. This might entail a lack of accounting for clinical “fine-tuning”. This touches on the very nature of clinical judgement and experience. Webster and Ben-Aron (1985) pointed out that: “Clinicians must surely use themselves as the microscope” (p. 50). It takes training and persistence to learn to use the microscope. At first the untutored eye only sees blobs of matter. Only gradually does one learn to see specific entities and to name meaningful structures.

It is important to point out that Ennis and Litwack (1974) found clinicians no better than chance *on average*. This means that there may well be clinicians who are “good with the microscope”, that were able to use their clinical judgement to predict violent outcome. Unfortunately, other clinicians operated on a level of chance, “flipping coins”, or worse, explaining the *average* lack of predictive validity. The reductionism approach to risk assessment consequently runs the risk of disposing of the “baby with the bathwater”. In my opinion, an alternative approach might be to study those clinicians that are able to predict better than chance in their clinical judgements and to study the nature of their assessment process to elucidate successful strategies in risk assessment.

Another problem in risk prediction concerns the way in which “predictive power” or accuracy is evaluated. It typically measures outcome in a dichotomous

(“yes” or “no”) fashion. This does not allow for differences in magnitude or imminence of the violent outcome, both of which are factors that are crucial to the identification of risk. Moreover, this model for predictive accuracy has a more practical than clinical meaning: “Dangerousness should be conceived rather as a continuum of potentiality” (Menzies, Webster, & Sepejak, 1985, p. 51). The dichotomous predictive validity is typically described in a 2 x 2 table (adapted from Monahan 1981, p 47): See Table 1.

**Table 1.** Evaluating predictive accuracy.

		Actual Outcome	
		Violent	Not Violent
Predicted Outcome	Violent	True positives a	False positives b
	Not Violent	False negatives c	True negatives d

This model puts demands on deciding specificity and sensitivity (Yerushalmy, 1947) of risk assessments. The choice is not in the predictive validity domain of risk assessment, but rather in the risk decision domain. Given that the predictive validity so far is not 100% correct, we need to decide on the nature of the false decisions we are prepared to make, in terms of either accepting a number of false negatives (c) or false positives. (b). The overall correct prediction is described by  $(a + d)/N$ . The positive predictive value (PPV) is  $a/(a + c)$ , and the negative predictive value (NPV) is  $d/(c + d)$ . The specificity of any risk assessment procedure is proportion of actually non-violent cases, which are correctly predicted,  $d/(b + a)$ . The sensitivity of any risk assessment is the ability to identify violent recidivists or the proportion of actually violent persons who are correctly predicted:  $a/(a + c)$ . A commonly used methodology to handle this problem is to use Receiver Operating Characteristic (ROC) analysis, a statistical procedure, which requires a dichotomous outcome and depicts the trade-off in specificity relative to sensitivity, described in a curve for an optimal balance. The overall accuracy in this measure is the expressed as the Area Under the Curve (AUC). Originally developed for radar signals detection and used in radiology (Hanley & McNeil, 1982), this analysis has come to be a more or less standard way of reporting violence prediction validity in (Douglas et al., 1999; Douglas et al., 2003; Grann et al., 2000).

### *Predict or Prevent?*

A perhaps more serious problem than the somewhat artificial debate on actuarial versus clinical assessments that have been presented from the “second generation” of research has been the emphasis on prediction. The use of models for risk assessment, whether actuarial or structured professional judgement-based, that reduces the complexity of violent behaviour into 12 or 20 items is of course problematic. Violence is best conceptualised as a multi-faced social, psychological and biological human characteristic and it is naïve to think that we ever will be able to fully predict its occurrence. Despite the need to be modest about our efforts to do so, we are obliged to try. John Monahan (1984) was instrumental in this development in stating that:

“ What little we know can be improved upon; and how useful this knowledge is depends upon what we do with it, **compared with what we would do without it**” (p.13, emphasis added).

Static risk factors are often of limited utility in clinical practice. Historical data may give a fair estimate of the level of risk, but is rarely useful for the management of risk. Static, or historical risk assessment can only be meaningfully used once and cannot inform as to the risk is increasing or decreasing over the years. If one tries to use this type of risk assessment as a repeated measure, it can become quite paradoxical. Using the Violence Risk Appraisal Guide (VRAG; Quinsey et al., 1998) this becomes apparent. A patient with a personality disorder, who previously murdered a male victim may lower his risk considerably by obtaining a diagnosis of schizophrenia, marry and then kill his wife. Undoubtedly, very few clinicians would recommend this form of risk management! The risk predictions also assumes that no interventions are made to manage the risk. In most walks of mental health practice, *predicting* behaviour is not the task. Swedish legislation on forensic psychiatry (LRV §16) on discharge of patients and the Law of correctional facilities (KvaL § 32) on unsupervised leaves, respectively, clearly states that it is the obligation of clinicians and decision-makers to prevent acts of violence, not merely predict them. Most clinicians also view prediction only as professionally dubious. Dr. Stephen Hart argued that

“Clinicians are bound, morally, ethically and legally, ...to prove themselves wrong when they ‘predict’ violence; they must take every reasonable action to ensure that those at high risk for violence do not act violently” (Hart, 1998, p 123).

### **DEFINING RISK ASSESSMENT**

The research questions are focused on the accuracy and process of risk assessment, in the context of the practicing clinician, including the decisions made concerning risk.

To be noted is that the term of risk assessment, in this context, refers to the risk of repeated offending, recidivism, not to the risk of the first violent offence either in patients or members of the general public.

Definitions of risk describe a state of uncertainty about the future. The concept of risk typically requires at least three factors. The first is that the concept of risk is used for appraising undesirable, aversive events, such as violence, accidents or monetary losses (for desirable events the term is labelled chance). The second factor is the appraisal of probabilities of the occurrence of the aversive event or events. Finally risk also considers the consequences of the event occurring. Typically the acceptance of risk is a balance between the latter two factors. If the consequences are very severe the probability must be appraised to be extremely low for the risk to be acceptable (as with the risk of accidents in a nuclear power plant). In the opposite direction, if the consequence is judged as very light or unimportant people might be willing to accept a probability exceeding 99% (as in losing 50SEK in the national lottery).

Risk assessment, as it is defined in the present thesis, is broader than risk prediction. Although papers 2 and 4 do measure actual violent outcome, the main interest in all 5 papers is with the process of risk assessment as it is applied in the ecological, clinical practice. The definition used in this thesis adheres to Heilbrun's definition (2001; 2004). This definition outlines the stages or elements of risk assessment:

1. Identification of risk (risk of what, to whom, when under what conditions?)
2. Risk prediction (what is the probability of violence, identified above?)
3. Risk Management (How can the event be prevented by intervention or monitoring?)
4. Risk Communication (how can the risk be understood in a clear and unambiguous way?)
5. Risk decision-making (decisions related to violence risk, detainment, commitment, parole, release etc).

In earlier studies of this area, the concept of “dangerousness” was typically used (Webster & Ben-Aron, 1985). However, this concept is “unavoidably vague” (Monahan, 1981).

Monahan argued that dangerousness confuses what is being predicted with the probability of the event occurring. “ /--/ ...the word has a tendency in practice to degenerate from a characteristic of behaviour to a reified personality trait” (p. 25). Monahan went on to argue that “prediction of dangerous behaviour” is no less ambiguous. This concept is a conditional probability. If one steps on dangerously thin ice, then one will fall in the water. Litwack concluded that a clinical assessment of dangerousness is not equivalent to a prediction of violence” (2001, p. 172).

For the purpose of this work I have chosen a definition of risk assessment suggested by Hart; “*Risk assessment is the process of understanding a hazard with the aim to limit negative impact*” (Hart, 1998). It is a broad definition and avoids the risk of overemphasising probability in comparison with other factors, particularly the consequences of the event occurring. Recent studies on the accuracy of risk assessment focus on accuracy of probabilities, but pays little attention to the magnitude (consequence) of the violent crimes predicted. For outcome purposes reconviction of a violent crime is often used. However, most people would consider that there is a great difference between violent crimes such as murder and minor assault. Both would however be considered violent recidivism and calculated with in the fit of the prediction model studied (Belfrage, 1998; Belfrage & Douglas, 2002;

Douglas et al., 1999; Grann et al., 2000; Müller-Isberner et al., 1998; Tengström, 2001).

## **ETHICAL CONSIDERATIONS IN RISK PREDICTION**

The ethical issues pertaining to risk assessments are complex. At the core of the ethical balance of risk assessment practice are the potentially different best interests and perspectives of the involved parties. There is a potential conflict between civil and legal rights of the person who is assessed and the need for public protection. The rights of the prisoner/patient may at times be infringed upon in the interest of society and its agencies (courts, tribunals) with vested legal powers to detain and incarcerate individuals. These agencies are obliged to protect citizens from potential harm. In this process the legal agencies have the role of acting on behalf of potential future victims of violent crime. At the heart of the matter, the process of risk assessment implies that the propensity for future violence is measured and that decisions potentially infringing on personal freedom are to be made proactively, i.e. individuals are sometimes indefinitely committed for acts that they have not yet committed, and maybe never will commit. Jurisdictions differ in their approach. Risk assessments make the difference between reaching the parole-board or not in many U.S. states. In Britain, the “Dangerous and Severe Personality Disorder Program”, piloted by the Home Office, has been criticised for proactive detainment of offenders (Buchanan & Leese, 2001).

The tasks of public protection, securing legal rights of patients, and offender rehabilitation might place demands on the clinician that are incompatible and cause role conflicts for the practicing clinician. Appelbaum (1997) argued that clinical and forensic roles are incompatible and mental health professionals should decline risk assessments for their clients. On a practical level, the conflicting demands on the clinician can be resolved by assigning the tasks of therapist/helper and risk assessor to different clinicians. In many settings, struggling with crowded outpatient clinics, small forensic units etc., this can, however, be hard to achieve. The forensic practice of “expert” clinicians who make risk assessments away from the clinical context of the patient have ethical advantages, but it might also reduce the role of the clinician to “probability technician”. This might reduce the ability of the clinician to put focus on risk management and risk-reducing interventions.

The Hippocratic oath (cure, alleviate and comfort) is the guiding ethical principle of physicians and other clinicians working in medical contexts. This guidance does not specifically address the protection of others or of society at large. There is an absolute obligation of the clinician to act in the best interest and not to cause harm to the patient. In clinical psychology and psychiatry the role of the clinician is typically that of a therapist. The relationship to the patient is safeguarded by confidentiality and careful documentation in the clinical context (APA, 1992). Furthermore, the working alliance between therapist and patient is one of the key factors for successful treatment (Weinberger, 1995; Seligman, 1995). In many cases, however, there is little or no controversy in combining the roles of clinician and risk assessor. It can be argued that it is not in the best interest of the patient to be put at risk of committing further violent crimes. However, few patients see themselves as presenting a high risk of violent

crime, and the onus is potentially very strong on the nature and credibility of the risk assessment.

The “who should do it?” question also raises ethical issues about golden standards of assessment and assessor training. Furthermore, the issue of accountability of the risk assessor will vary in different legal frameworks. Even in settings where the role of the assessor clearly is that of the expert witness and the risk decisions rest with the court, or equivalent, there are still moral and ethical obligations of the assessor (Monahan, 1984). To keep the practice of risk assessment safe from a patient perspective there needs to be some requisites. There needs to be guidelines in the legal context as to who should be assessed and when the assessment should be done. There is an obvious risk that risk assessments are not kept up-to-date. There also needs to be guidelines for risk acceptance.

What is considered enough risk for intervention? In a recent study, Monahan and Silver (2003) found that some Judges were not prepared to accept any risk of violent reoffending at all in mentally disordered offenders. There also needs to be a transparency in the risk assessment that lends itself to second opinions and to the challenge of procedures and results. Although risk assessment is ethically problematic, legal decision-making requires that evaluations of risk of violent reoffending be done. It is then an ethical practice to do this in accordance with scientific knowledge and in an ethical framework (Grann & Nilsson, 2000). To claim that it should not be done, on ethical grounds, leaves the field open to the unscientific speculations of unstructured clinical judgement (Nilsson, 2000). This would be both deeply un-ethical and nonsensical.

## **METHODS OF RISK ASSESSMENT**

### **Actuarial versus structured clinical methods**

The need to assess risk of serious crimes to health and safety of others in certain individuals does not disappear because of ethical considerations of the professionals. Even against the pessimism of the 1970's, legislators and the general public still felt the need of guidance in decisions of incarceration and other measures to some individuals on the basis of their danger to the community (Shah, 1978).

As a reaction to the arbitrary nature of clinical judgement, researchers argued for a second generation of risk assessment techniques, namely the actuarial methods of risk assessment (Menzies & Webster 1989; Monahan, 1984). The term of dangerousness was considered outdated because of the connotations that this concept has with risk of violence as a dichotomous trait, possessed by some individuals. The term of risk assessment was thought to be more dynamic and reflecting a view of ‘risk’ as a continuous variable, that also might change over time and contexts.

Actuarial assessments do not take place inside an expert's head (Meehl, 1957), and the process of assessment is made from clear criteria and described variables, which do not require clinical expertise. Actuarial assessments could be argued to be more transparent and open to scrutiny than clinical judgement, which is based on an individuals experience and opinions. The great advantage of actuarial methods was that it is less arbitrary and the validity of the methods were open to study and

improvement by empirical research on what kind of variables actually are empirically associated with recidivism in violent crime.

Actuarial methods are sometimes called mechanical or statistical prediction models. Used as a model to collect, systematise and analyse data to guide decision-making were introduced into violence risk assessment in the 1980s (Monahan, 1984, Webster et al., 1994). They were, actually, demonstrated to be superior decision-making tools in clinical psychology and psychiatry, decades earlier, compared to unstructured clinical judgement or “gut feeling”, often used in the clinical practice then and perhaps still (Meehl, 1957). In recent years, a number of actuarial prediction models for violence risk have been suggested and tested in the literature on violence risk assessment. Many of these originate from North America, but trials have been conducted in Europe as well. Some of the more successful actuarial tools are *Violence Risk Appraisal Guide* for violence (Quincey et al., 1998) and the *Static-99* for sexual violence (Thornton & Hanson, 1999).

It is possible to perceive the controversy between actuarial and structured clinical assessments from an either/or perspective. This perspective is probably not fruitful and Webster and colleagues (2002) has successfully argued that this debate has outplayed its best by date. There is evidence for the predictive validity of both approaches. Hanson (1998) suggested that that actuarial assessment of sexual recidivism would eventually outperform clinical judgement as more knowledge is amassed. However, Litwack (2001) found that there have actually been few studies that compares actuarial versus clinical methods in the evaluation of recidivism risk, using the same subjects population and similar outcome criteria. Furthermore, the concept of performance in this context refers only to the predictive validity. This has a number of limitations in the view of the definition of risk assessment described above (Heilbrun, 2004).

There are sometimes misunderstandings about the nature of actuarial assessments. Actuarial refers to the method of gathering and combining data; for a thorough description see Sawyer (1996). Because of the statistical nature of actuarial assessment and that this type of assessments often draws on historical data, it can be confused with static assessments and some risk variables are sometimes referred to as actuarial items (Harris & Rice, 1997). In my opinion, however, this is not a correct use of the term actuarial. Items or risk factors can be static or dynamic and the method of their assessment can be actuarial or based on clinical judgement (whether impressionistic or adhering to guidelines). See table 2.

**Table 2.** Examples of assessment approach and stable-dynamic factors

<b>Assessment Approach</b>	<b>Changeable</b>			
		<i>No Static</i>	<i>Yes Stable Dynamic</i>	<i>Yes Acute Dynamic</i>
	Actuarial	Earlier Violence Index Crime	Age Ever Married	Sero-positive alcohol Recent divorce
Structured Clinical	Brain damage Autism Learning difficulties	Insight Attitude Personality Disorder	Responsivity Intoxication Emotional state	

Litwak (2001) concluded that actuarial and clinical assessments are different in nature and that there in fact have been few studies directly comparing clinical versus actuarial assessments of violence risk using the same subject population and outcome criteria.

**Developmental basis of risk assessment methods**

In order to be able to prevent violent acts, the emphasis must rest on dynamic risk factors to amass knowledge on effective risk management. The current knowledge of evidence-based risk factors makes the prospect of doing this more feasible today than a few decades back, when no valid risk assessment procedures were available.

The nature of the samples upon which the inferences on probabilities are drawn may be of interest. Risk assessment tools are essentially developed in four different ways, or in combination of these. Table 3 is a description of these methods and gives examples of risk assessment procedures developed by the methods respectively.

**Table 3.** Developmental procedures of risk assessment tools.

<b>Method</b>	<b>Example</b>
<b>Sample-based</b>	
1. Single sample	Violence Risk Appraisal Guide (VRAG) (Quinsey et al., 1998)
2. Multi-site sample	MacArthur Tree (Steadman et al., 2000)
<b>General literature based</b>	HCR-20 (Webster et al., 1997) SARA (Kropp et al., 1995)
<b>Meta analysis based</b>	Static 99, Sonar 2000 (Hanson & Harries, 1998)
<b>Serendipity based</b>	PCL-R (Hare 1991)

A typical example of a single sample approach is the VRAG. This came from years of studying patients released from one institution, Oak Ridge in Ontario, Canada (Quinsey et al., 1998).

The chosen risk factors could be obtained and measured from patients' life stories as they appeared on file and records. For economical and practical reasons the data is typically gathered retrospectively in this kind of study. They also rest on the assumption that a certain combination of these factors "predicts" recidivism, although the results only suggests a correlation between a risk factor and the outcome criterion, which is typically further arrests or verdicts for violent behaviour. This approach have been criticised because of these methodological problems (Grann & Långström, in press; Nussbaum, 2004). The criticism has focused on the clinical usefulness of this kind of assessment. A single sample design is also frail because of the assumption is that any individual who is assessed beyond the original sample is representative of the same sample. The prediction algorithms in this kind of method are optimised to the original sample, and any further samples (which is not identical with the original) will have a less good "fit" to the algorithm. It also expresses absolute probabilities for the individual case (46% risk in 7 years), which gives a flare of absolute "scientific

exactness” about the prediction. Most researchers will agree that risk prediction is not rocket science.

A related approach is derived from a developmental sample from multiple sites. This might avoid some of the frailty of a single sample, but is essentially comparable.

The serendipity of the PCL-R simply refers to the fact that it was not originally developed as a risk assessment tool (Hare 1993), but later proved to be highly associated with violent recidivism (Harris et al., 1993; Steadman et al., 2000).

In this respect the “general literature” and the meta-analysis approaches might be more robust. This is very much the successful “Vancouver approach” resulting in a number of risk assessment procedures for different groups of offenders from the research group with professor Christopher Webster and co-workers.

### **Structured Clinical Methods**

Structured clinical assessments of risk have sometimes been referred to as the “third generation” of risk assessment (Grann, 2002). The first generation being the unaided, unstructured clinical judgements or clinical opinions (Dolan & Doyle, 2000), and the second generation being the actuarial risk prediction algorithms called for by Monahan (1984). The structured clinical approach is based on principles of “anchored clinical judgement” (Quinsey et al., 1995). It relies on an actuarial method on judging the relative value of historical, static risk factors for reoffending. The approach also incorporates clinical judgement, but this judgement is guided to violence-relevant issues. This method takes on the form of an “aide-memoire” or “assessment protocol” for assessing risk of offenders that draws on findings in the risk assessment literature (Blackburn, 2000). The virtues of this rather pragmatic approach is that it requires the assessor to attend to information or items that might otherwise be overlooked and in the detailed descriptions of the risk factors to be looked for by the assessor or clinician. The clinician may effectively take full advantage of his or her knowledge of the unique characteristics of the case. Another prominent feature of the approach is the focus on risk management. These methods tend to underline dynamic risk factors for their usefulness in the clinical practice. This approach is usually referred to as structured clinical assessment, but sometimes called “professional clinical assessment” or “empirically guided risk assessment” (Hanson 2003).

A common factor of these methods is that they originate from Canada. The “Level of Service Inventory” (LSI) was developed in Canada by Andrews and his colleagues (Andrews & Bonta, 1995) to assess risk and needs of offenders (predominantly without mental disorder). There has been substantial support for validity and usefulness in the literature (Andrews & Bonta, 1998; Palmer, 2001). More importantly for the present thesis is the work by Webster and colleagues at Simon Fraser University in Vancouver, British Columbia. This group has produced an array of risk assessment guides for different purposes. The common characteristics of these guidelines are that they are based on the scientific literature on risk assessment and risk factors and that they use the assessment technique of the Psychopathy Checklist; PCL-R (Hare 1991; 2003). This entails asking clinicians to rate carefully described items in

categories of 0 (does not apply), 1 (applies to some extent) and 2 (definitely applies). These tools also keep to a format of around 20 items. This is not for any particular scientific reason, but it is a manageable number of factors and the added information value just may not increase exponentially with further items added (Webster et al., 1997). Methods using the “Vancouver approach” are: *The Sexual Violence Risk, SVR-20* (Boer et al., 1997), *The Spousal Assault Risk Assessment, SARA* (Kropp et al., 1995), *The Early Assessment of Risk List- Boys, EARL-B* (Augimieri et al., 2001) and *The Historical-Clinical-Risk assessment scheme, HCR-20* (Webster & Eaves, 1995; Webster et al., 1997). For an in-depth overview see Douglas, Cox, and Webster (1999). The HCR-20 comprises, unsurprisingly, 20 items, all rated in the 0-2 fashion of the PCL-R (Hare, 1991; 2003). The assessment protocol is organised in three sections, Historical Items, Clinical Items and Risk Management (situational) items. At the end of the assessment the clinician is asked to do a final risk judgement of low, moderate or high risk.

Historical items are all static in nature (by definition they have already happened). The clinical items are dynamic and refer to the current condition of the person assessed. The Risk management items are also dynamic (by default, because they haven't yet happened) and focuses, not on the individual, but on a review of the treatment and management plans for the individual. For an overview see table 4, next page.

**Table 4.** The HCR-20 assessment scheme

<b>Historical Scale</b>	
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
<b>Clinical Scale</b>	
C1	Lack of Insight
C2	Negative Attitudes
C3	Active Symptoms of Major Mental Illness
C4	Impulsivity
C5	Unresponsive to Treatment
<b>Risk Management Scale</b>	
R1	Plans Lack Feasibility
R2	Exposure to Destabilisers
R3	Lack of Personal Support
R4	Non-compliance with Remediation Attempts
R5	Stress

The HCR-20 scheme has also had the advantage of being the subject for numerous studies and evaluations both in different legal contexts, i.e. correctional, forensic psychiatric and civil psychiatric and in different countries (e.g.: Belfrage, 1998; Belfrage & Douglas, 2002; Dolan & Khawaja, 2003; Douglas et al., 1999; Grann et al., 2000; Müller-Isberner et al., 1998; Philipse et al., 2000; Tengström, 2001). The assessment scheme was supplemented by a guidebook, focusing on risk management in 2001 (Douglas et al., 2001). Kevin Douglas also keeps an updated annotated bibliography of studies on the HCR-20 from all over the world ([www.sfu.ca/psychology](http://www.sfu.ca/psychology)).

## **IMPLEMENTATION INTO CLINICAL PRACTICE**

Choosing a valid, reliable and practical risk assessment tool is but the first step in implementing risk assessment procedures into clinical practice. The next step is to study how clinicians and other assessors use the procedures and what influences this process. Findings from clinical psychology research on clinical practice could provide

some guidance. Previous reviews have been quite pessimistic about clinicians' ability to make decisions on diagnosis, case-formulation and prognosis of treatment success (Gough, 1963; Kleinmuntz, 1984; Dawes, 1994).

## **Biases and threats to the process of risk assessment**

### *Expertise problems*

A potential problem in clinical risk assessment may be that clinicians are just that; clinicians. The criterion variable, future violent crime, is not necessarily a focus for clinicians and indeed not something they are trained to assess or manage. The assumption of clinical expertise in this field rests on the notion of a link between mental disorder and violence (Ennis & Litwack, 1974). This assumption is unsafe and the relationship is quite complex (see section 2, Violence and mental disorder). The omnipotent expectations on clinicians are not a new phenomenon. Holt (1970) reviewed the literature and found that clinicians often were asked to predict things like academic achievement or violent behaviour, for which they had no expertise. Rock et al. (1987) argued that research on the lack of validity of unstructured clinical assessment did not take into account what types of assessments and in what contexts clinicians are trained to do assessments. They argued that this research was essentially non-ecological. Clinicians can of course also be misinformed or just plain wrong. If risk factors like epilepsy or enuresis is considered in risk assessment the validity will inevitably be low (Prins, 1986). Clinicians are trained to diagnose mental disorders, conditions and psychological disturbances. Although some conditions are associated with violence (personality disorders) others are not (sleep disorders). The taxonomic systems (DSM & ICD), used in mental health services are aimed at bringing order into diagnosing and to assure some reliability in psychiatry. They are not aimed at prognosis or even aetiology and have a limited contribution to the understanding of violent behaviour (Murphy & Clare, 1995).

### *Information processing and clinical decision-making.*

Clinical assessments and risk assessments are no exceptions to the limitations of human information processing and decision-making in general. The limited rationality is a fundamental issue in human reasoning (Kahneman & Tversky, 1996). Studies have shown that people tend to overemphasise some cues to generate hypotheses or evaluate conditions generally. This might be a model for understanding why retrospective, actuarial assessments typically have better validity than unstructured clinical interviews. Tversky and Kahneman (1974) identified some of these cues and "shortcuts" in human reasoning. A "shortcut" or heuristic was labelled "anchoring". This describes a tendency to stick with the initial assessment and structure further information to confirm this assessment. The so-called illusory correlation bias (Chapman, 1967) describes the tendency to attribute association and causal relationships in uncorrelated events or to overrate weak associations. A related problem is the limitation in handling large amounts of information (Leuger & Petzel, 1979).

Clinical judgement and risk assessment are based on extensive and complex information from a number of different sources (observations, interviews, tests, references, files and documents etc.). The complexity of clinical case formulation (Phares & Trull, 1997) makes actuarial methods hard to use, or at least requiring information reduction that clinicians are sometimes unwilling to do (Boothby & Clements, 2000). Almost 50 years ago Paul Meehl stated that “*mostly we will use our heads, because there just aren't any formulas*” (1957, p 237), when describing clinical decision-making. To further dampen the prospects of clinical assessment there is ample evidence to support that experience and training generally does not improve the validity of clinical assessments (Garb, 1992; Dawes, 1994).

### *Irrational influences*

The practitioner in clinical practice is primarily interested in understanding and building models for the thoughts, emotions and behaviour of the patient. The main objective of the clinician is to influence the patient in a favourable direction concerning thoughts, feelings and behaviour. It is sometimes inadvertent that this process goes the other way and the patient influences the clinician. In traditional psychoanalytical practice, this is referred to as counter-transference (e.g. Armelius & Holmqvist, 1996). These feelings in the clinician may be irrational or at least irrelevant to an assessment context (Cohen, 1983).

So far there have been no empirical studies of irrational influences on structured clinical risk assessments. However, there is a wealth of empirical studies on the effects of therapists' feelings in the treatment literature (for an overview see: Cullari, 1998).

## **RISK COMMUNICATION**

The communication of risk of violence has recently been considered an important part of risk assessment process (Heilbrun et al., 1999). Some empirical studies have shown that communication style does influence risk decisions. Probabilistic versus categorical risk statements might change the perception of risk of the user (Murphy & Clare, 1995). Even when using probabilistic reporting, the scale of estimates (Slovic & Monahan, 1995) and expression of risk in probability (20%) versus frequency (1 in 5) formats influences the perception of risk (Slovic et al., 2000; Harries & Harvey, 2000). Recently, Monahan and colleagues (2002) found that expressing probability in frequency formats led to more conservative discharge decisions than probability formats.

There are some suggestions in the international literature on which issues should typically be addressed in reasonable risk communication about violence risk (Heilbrun et al., 1999; Slovic, 2000). These include a statement about the level of risk (e.g. very low, low, medium, high or very high), the nature of the risk, factors that increase or decrease the risk (see also Elbogen et al., 2002; Sturidsson et al., 2004), the imminence or which timeframe the risk statement apply to, in which situations, and finally suggestions for risk management and treatment. Grann and Pallvik (2002) studied the written risk communication of 142 cases of forensic psychiatric evaluations in Sweden. They developed a study protocol for determining the level of elaboration in the risk communication using six criteria from adapted from Heilbrun et al (1999), and they reported that risk was well elaborated only in 17% of the reports.

## **RISK DECISION-MAKING**

There are, at least two important aspects of decision-making pertaining to risk generally, not exclusive to violence risk decisions. The first aspect is connected to the communication of risk and could be conceptualised as risk-perception. Research has shown that the perception of risk is not rational (Slovic, 2000). The perception of risk is influenced by a tendency to overweigh small probabilities relative to medium or large ones, as described by Kahneman and Tversky (1979). An example of this is that in some situations people will use a high-probability term such as “good chance” to describe numerically low probabilities (Teigen, 2001). The perception of risk is also greatly influenced by the perceptions of other individuals and groups of individuals (Davis, 1992). Another important aspect of risk-decisions could be conceptualised as risk acceptance.

Behavioural Decision Research points to a number of factors that might influence risk acceptance in decision-makers (Baron, 1994). Swets and colleagues (2000) pointed out that judges need to apply some form of “decision-thresholds” or cut-off for what degree of risk should be accepted. Silver and Monahan (2003) found that 12% of Judges in criminal courts had a decision threshold of 0, i.e. they accepted no risk at all in mentally disordered offenders. A majority would not accept a risk of more than 26% and 23% would reach their acceptance threshold when the risk exceeded more likely than not to re-offend (i.e. >50%). The authors also point out that the special case of making decisions about violence risk management on the basis of risk analysis seems to be understudied, in fact, the Authors claim theirs to be the only study of risk-acceptance thus far. This is perhaps surprising since decisions about, and appraising other types of risk, particularly financial risk, is a key-area in the psychological literature (Conolly et al., 2000). There are a number of theories that might be appropriate for the study of this area, such as evolutionary game theory, bounded rationality or ecological rationality (Gigerenzer & Goldstein, 1996)

In this context, the Nobel Prize awarded theory by Kahneman and Tversky (1979), seemed of particular interest. This is named Prospect theory and views outcome as gains or losses from the subjective value function. Very simplified, Prospect Theory provides a framework for understanding how the perceived gains and losses might govern decision-making. It also gives some guidance into how heuristics and biases might be involved in this process (Kahneman, Slovic, & Tversky, 1982). Very simplified, prospect theory states that people are cautious about obtaining gains, preferring “sure-things” to gambles, while being risk-seeking about potential losses; to avoid a certain loss people will take a gamble that could lead to an even bigger loss. Prospect Theory is also a leading theory of decision-making, and it has been demonstrated to describe and predict a wide range of data (Hardman & Harries, 2002). It was particularly the subjective “value function”, the basis for the theory that made it interesting to violence risk decisions. This function is able to predict ‘framing effects’ that might be important in decisions about future prospects of violence.

For the purpose of the present thesis, I had an interest to assess how the potential losses (negative consequences) and perceived gain (positive consequences) would be valued by decision-makers for the issue of releasing a mentally disordered violent offender (posing a certain risk), into the community. It was particularly interesting to examine differences in subjective value function between groups who are exposed to this kind of decision-making in their professional capacity on an everyday basis such as criminal law professionals and clinicians compared to a control group, who do not.



## **AIMS**

The ultimate purpose of the present thesis was to advance knowledge regarding how to improve risk assessment in order to decrease the risk of violent recidivism in mentally disordered offenders.

The overall aim of this thesis was to study structured clinical risk assessment procedures as they are applied in forensic psychiatric practice and how it is influenced by technique (HCR-20), assessor characteristics, risk management and influences risk acceptance.

The specific aims were to:

- i. To empirically study the inter-rater concordance and the content validity of the HCR-20 as it is applied in a clinical setting
- ii. To empirically study reciprocal effects of risk assessment and the level of risk management
- iii. To empirically study irrational factors that might influence the structured clinical risk assessment procedure.
- iv. To study the role of how information is used and valued and to elucidate the effects this might have on structured clinical risk assessment.
- v. To elucidate whether risk assessor training, position to the assessed patient had effects on violence risk prediction in the short and long-term.
- vi. To empirically study the effect on risk acceptance and risk decision-making of factors from human decision making theory.



## **METHODS**

### **SETTING**

The studies in the present thesis were all undertaken in the “risk program” financed by the County Council of Östergötland. This program was initiated in 1997 at, what was then, Rättspsykiatriska Regionkliniken, RPR (the regional forensic psychiatric unit) has been running at, what is now, RättsPsykiatriskt Center, RPC (forensic psychiatric Centre) in Vadstena Sweden. This program had a clear aim at improving the standards of risk assessment standards of patients at the Centre. The programme stems from a forensic practice need and from a clinical perspective. The program consequently included research questions about the risk and management of patients as well as the effectiveness and implementation of clinical risk assessment and management procedures. From 1997 to 2003, the formal academic link for the program was with the department of Psychology and Education (renamed dept. of Behavioural Science in 1999). From February 2003, the formal academic link was established to the Centre for Violence Prevention (CVP) at Karolinska Institutet in Stockholm.

### **Forensic Psychiatry in Sweden**

Studies I, - IV all were conducted with patients at the Centre, who were convicted to involuntary forensic care, according to Lagen om Rättspsykiatrisk Vård (LRV, 1991:1129), the Swedish legislation on offenders, suffering from “serious mental disorder”. In the Swedish legal system, the criminal courts that suspects an offender to be mentally disordered can refer the offender to a forensic psychiatric evaluation (FPE). The referral decision is at the discretion of the court. In the FPE, it is decided whether the offender suffered from a “serious mental disorder” at the time of the offence. If this is the case, the offender must not go down the correctional system i.e. prison. The FPE also decides if the serious mental disorder is still present at the time of the evaluation. If so, the common pathway is for the court to pass a sentence to forensic psychiatric inpatient treatment on non-fixed terms. At the FPE the risk of re-offending is also assessed. If this risk is deemed high, the sentence is typically one of “Särskild Utskrivnings Prövning (SUP), “special discharge procedures”. In these conditions decisions about leaves and discharge cannot be made by the hospital, but must be decided in a civil court (länsrätten). The Swedish legislation has, at the moment, no equivalent of “unfit to plead”, “diminished responsibility” or “not guilty by reasons of insanity” of Anglo-Saxon law tradition. The medico-legally insane are convicted and considered responsible (see Dernevik 2002).

## **SUBJECTS**

Studies I - IV were all conducted with patients at the Centre in Vadstena, who volunteered to take part in the research and were convicted of violent crimes, to inpatient forensic treatment and subjects to “special discharge procedures”.

Study I included 6 professionals who assessed 6 male patients for risk of violent re-offending.

Study II, included 54 consecutive patients, who were assessed following admission to the Centre. 48 were men and 6 were women.

Studies III and IV, included 8 volunteering patients each. 4 patients were the same individuals in both Studies. In study III all patients were men, and in study IV, 7 were men and 1 was a woman. These studies also included 40 volunteering members of nursing staff per study. Despite the equal numbers, they were different samples and the data was collected during staff training on two different occasions. One ”expert” risk assessor also assessed the patients. In study III, the expert was one out of five and in study IV, one out of six.

Study V, did not include patients. Instead, 56 criminal law professionals and 56 clinicians were asked to participate. They had all attended a practitioners’ conference, arranged by the Centre. A control group of 56 high school teachers, who did not attend the conference, were also asked to participate. 36 criminal law professionals agreed to participate, 32 of them completed the questionnaires and were included. 35 clinicians agreed to participate, 30 was included (same reason). In the control group 28 agreed to participate, 26 were included. In total 88 subjects were included in the study.

**Table 5.** Subjects in Study I - V.

<i>Study</i>	<i>Subjects</i>	<i>Points of observation (n)</i>
<b>I</b>	Clinicians rating forensic patients In complete rotation	6 x 6
<b>II</b>	Consecutive patients convicted of violent crimes admitted to forensic psychiatric hospital	54
<b>III</b>	Members of staff rating (8) well known, forensic patients, 5 raters/patient. Compared to 1 of 5 expert assessors	40 + 5 (8)
<b>IV</b>	Members of staff rating (8) well-known forensic patients, 5 raters/patient. Compared to 1 of 6 expert assessors	40 + 6 (8)
<b>V</b>	3 groups of decision-makers, Criminal Law Professionals, Clinicians & Controls	88

## **DESIGNS AND PROCEDURES**

Study I, used a 6 x 6 design. 6 clinicians all assessed the same 6 forensic patients, using the HCR-20 risk assessment protocol. Coherence between raters was compared. Content validity was estimated using a triangulation procedure: HCR-19 to Previous violent crime to future violent crime. One item from the HCR-20 scheme (previous violence) was excluded and HCR-19 scores were correlated with “Crime Index”, a measure of previous violence. The association between previous and future violence was estimated from a previously published study by Teplin and colleagues (1993).

Study II, used a prospective, cohort follow-up design. 54 consecutive mentally disordered offenders were risk assessed on admission to hospital by one of five experts, trained in risk assessment. The inclusion criteria were a conviction for a violent crime and suffering from a “severe mental disorder”. Subjects were monitored for violent incidents during in-patient treatment and follow-up for up to 52 weeks after conditional release and transfer to other facilities. The level risk-management was also monitored and classified as high, medium or low. The risk assessments were compared to the frequency of violence during the three management conditions.

In Study III, 8 patients were each, individually rated for violence risk by 5 members of nursing-staff, making 40 points of observation. The mean ratings for the 8 patients were compared to the ratings done by one of the same experts as in study II. The staff-raters were also asked to estimate their feelings towards their assessed patient. The ratings of feelings were compared to the variance in risk ratings.

Study IV, used a prospective, follow-up design for 8 patients (a partly different sample from Study III) The patients were rated for violence risk by 40 members of nursing staff (5 raters/patient). The patients were monitored for in-patient violence during hospitalisation and subsequently followed in the community for a further 48 months. Again the Nursing-staff risk ratings were compared to the regular expert ratings for the association with violent behaviour. Both staff-raters and expert-raters were asked for how useful they found four different types of information sources in the assessment process.

Study V, used an experimental 2 x 2 x 3 multifactorial design. Vignettes were used as stimuli. These were manipulated into high and low conditions for two independent variables, perceived loss and perceived gain, making a 2 x 2 = 4 conditions. Each subject rated 4 different vignettes with a varimax rotation of conditions, making a total of 352 units of observation (88 subjects x 4 judgements). This made analysis both within and between subjects possible. Three groups were compared on the outcome measure, “inclination to make discharge decisions”. Finally, a possible confounder, attitudes towards offenders, was controlled for using an attitudes questionnaire (ATMDO).

## MATERIALS

The Historical-Clinical- Risk assessment, **HCR-20** (Webster et al., 1997), was used as a risk assessment tool in all studies, except study V. The HCR-20 is described in detail in the section on methods of risk assessment (Structured clinical methods) in this thesis.

The concept of psychopathy is a fundamental part of risk assessment with the HCR-20. The construct is measured with the Psychopathy Checklist: Screening Version, **PCL: SV** (Hart, Cox, & Hare, 1995). It is a 12-item scale derived from the PCL-R (Hare 1991). It is scored in the same fashion as the PCL-R and the HCR-20, with the same 3-point scale (0-2) for all items. The maximum score is 24 and the recommended cut-off for psychopathy is 18 and higher (Hart & Hemphill, 2002). Like the PCL-R, the screening version in theory is made out from two factors; interpersonal/affective deficits and social deviance/impulsivity. Psychometric analyses indicate that PCL: SV has excellent structural properties and is strongly related to the PCL-R (Hart et al., 1995; Cooke et al., 1999). Cooke and Michie (2001) also found that the factor structure strongly paralleled that of PCL-R and that some of the 20 items in the PCL-R did not have high factor loadings in the psychopathy construct. The authors went on to suggest that the concept might be better captured with a 3-factors solution.

To compare HCR-20 scores with previous violence an independent measure, **Crime Index Scale**, was devised for Study I. The first item in the HCR-20 scheme is titled previous violence and this item was excluded from the analysis of correlation, making the “HCR-19”. The purpose of the crime index scale was to capture the extent and magnitude of previous violence. It was constructed as a 6-point scale. Item 20, criminal versatility from the PCL-R contributed with scores of 0-2. In addition to this the magnitude of previous violent crimes was rated 0-3 from criminal records. No previous violence scored 0 (not applicable in this context). Minor violent offences (resisting arrest) scored 1, assaults scored 2, and serious violence (murder, manslaughter) rendered a score of 3, thus making 0-5 points the total range of the Crime Index Scale.

To measure the assessors’ feelings towards the assessed patient (Study III), the Feeling Word Checklist (**FWL**) was used. Originally constructed by Whyte and colleagues (1982), it has been extensively used by Swedish researchers and the statistical properties have been described (Holmqvist & Armelius, 1994; 1996; 2000). In this study the scale was made out of 30 emotional adjectives, arranged in a circumplex model with four dimensions. These dimensions are: *Helpfulness* versus *Unhelpfulness*, *Closeness* versus *Distance*, *Accepting* versus *Rejecting*, and *Autonomous* versus *Controlled*. Technically, the adjectives are scored between 0-3 addressing the question, “when I think about NN, I feel:” The total word scores were added for each dimension and divided by the number of words in that dimension.

In study IV, a **usefulness attribution** was used to estimate the value placed on 4 different sources of information used while conducting an HCR-20 risk assessment for a particular patient. Subjects were simply asked to divide, out of an overall 100%, how useful they found:

1. The Reports from the Forensic Psychiatric Evaluation (FPE).
2. Hospital case-notes.
3. Interviews conducted for the risk assessment
4. Information from personal interaction with the patient.

In Study V, *Case Vignettes* were used as stimuli. The vignettes contained 4 different scenarios and 4 different offenders. This made a total of 16 mutations of vignettes. Each subject received 4 of these representing the 4 conditions (2 x 2), high gain/high loss, high gain/low loss, low gain/high loss, and low gain/low loss. The vignettes were short (250 words). They described offences that were topographically identical (bar fights), but the outcome was manipulated (manslaughter versus common assault) for the loss condition. The gain condition was manipulated by the future social circumstances of the offender (returning to work and family versus unemployment and hostel living arrangements). The probability of re-offending was kept constant (25%) in all vignettes.

Attitudes Towards Mentally Disordered Offenders (*ATMDO*) was controlled for in Study IV. Melvin, Gramling and Gardner (1985) constructed a 36-item questionnaire to measure general attitudes towards prisoners. Hogue (1995) adapted the ATP for the use of measuring attitudes to sexual offenders by changing the word “prisoner” to “sex offender” (ATS). For the present study we adapted the 36 items to mentally disordered offenders by replacing the word “prisoner” by “mentally disordered offender” (ATMDO). The questions were translated into Swedish by the author (MD) and the translated back into English by a colleague (Dr. Anne Frodi). Discrepancies between the original and the reverse translation were discussed and adjusted. The questionnaire uses a 1-5 points format making 36 points, the most negative (conservative) response possible and 180 points the most positive (liberal) towards mentally disordered offenders.

Finally, an *Outcome Questionnaire* was constructed for Study V. Six questions were asked to each vignette. They had a 6-point scale, ranging from definitely yes to definitely not.

The first and last of these were used as the dependent variable. Question 1 was designed to assess risk acceptance: *How inclined would you be to discharge the patient (if you were a member of the tribunal)?* Questions 2-5 were about risk of recidivism, how inclined other people would be to discharge, danger to others, if the subject can be rehabilitated into the community. They were not used in Study 5 and acted as “decoys” to make the objective of the study less obvious to subjects.

Question 6 was: *How long would you judge it appropriate for the subject to remain in a secure psychiatric facility?* Again a 6-point scale was used, ranging from “less than six months” to “for ever”.

## STATISTICS

The *power* of a study refers to the probability of detecting a relationship between variables or differences between groups. The power is dependent on the size of the study group (N), the magnitude of the effect size (ES), and the likelihood ( $\alpha$ ) to falsely detect a relationship or difference, where there actually is none. When designing studies it is essential to have a large enough sample to detect effect sizes (ES) of relevance for the research questions. In research performed close to the clinical practice, ecological settings, this can present problems (Rock et al., 1997). In study I, power was not an issue since this was a correlation study and low associations, even between few risk judges, would have entailed a rejection of assessment methods. In study II, the power was restricted by the cohort design of all admitted patients during 2 years. For 54 patients the power to detect a significant ( $\alpha < .05$ ) of correlations of .30 was 65%. In studies III and IV, the power was again restricted. The number of observations (40) for nurses meant that substantial effect sizes could have been missed. In study V, the power calculations pointed to group sizes of 56, to detect group differences at an acceptable level. The power calculations used EpiInfo free computer software (Centre for Disease Control and Prevention, 1994).

Pearson's product-moment correlation coefficients ( $r$ ) was used to elucidate relationships between risk raters in study I, to express associations between risk assessment and actual violence in study II, to explore associations between nurses' and experts' risk assessments in study III, to express associations between risk assessments and violent behaviour in study IV.

In study II, physical and verbal violence was analysed using partial correlations through the risk management conditions. This is considered a more stringent test than a simple correlations matrix. (Norman & Streiner, 2000). In theory the partial correlations between any two variables should be small after partialing out the effects of another, if there is an underlying factor structure.

Predictive validity in Study II was also studied by receiver operating characteristic analysis (ROC). This procedure has roots in signal detection theory (Hanley & McNeill, 1982). The ROC curve presents a graphical representation of the true positive rate (sensitivity) as a function of the false positive rate ( $1 - \text{specificity}$ ) across all possible cut-off scores of the assessment procedure. The positive and negative value of a test (PPV and NPV) is closely linked to the sensitivity and specificity, respectively. The ROC-curve can be used to explore how the specificity is influenced as the sensitivity increases. This is expressed as the Area Under the Curve (AUC). The total area under the ROC cut-offs curve can be seen as a measure of overall predictive accuracy. An AUC value of .50 describes the line of no information, while 1.0 indicates perfect predictions of dichotomous outcome (i.e. violent behaviour or no violent behaviour, see figure 1 in 3.2.1, violence prediction). There are no fixed evaluations but an AUC of .75 and above has been suggested as "large" predictive ability (Dolan & Doyle, 2000; Quinsey et al., 1998).

To explore the relationship between risk assessment and feelings towards the patient a two-way Analysis of Variance was conducted. A multiple regression analysis was also performed in study III. For study V, Chi-2 test was used for differences

between groups, t-test for differences within groups and ANOVA was chosen for the analysis of variance, because of there being more than 2 groups and multiple factors. The strength of relationships was drawn from the ANOVA in terms of eta-squared ( $\eta^2$ ). With the equation:  $\eta^2 = SS \text{ between} / SS \text{ total} = 1 - SS \text{ within} / SS \text{ total}$ , this will always yield a number between 0 and 1 and is interpreted as the proportion of the variance in the dependent variable (release decisions) that can be attributed to the independent variable (manipulated conditions) (Norman & Streiner, 2000).

The statistical analyses were carried out in the following programs: Study I, Statview for MacIntosh (Apple), Study III in SPSS for Apple 4.0. Studies II, IV and V were computed in SPSS (PC) versions 9.0 and 10.1.

## RESEARCH ETHICAL CONSIDERATIONS

Collecting and compiling data on personal background information and criminality is a potential threat to the personal integrity of the studied individuals. This also applies for the victims of the violent crimes who are described in the written and interview material that constitutes violence risk assessments. To protect the personal integrity of offenders and victims, sound research ethics needs to be adhered to. In the clinical context, offenders are not asked to agree to being assessed for risk, it is done with or without consent as part of the clinical procedure and by legal requirement. However, for research purposes consent would normally be needed. In Study I, III and IV information was given verbally and verbal informed consent was received from the patients.

In Study II, no informed consent was given. Because of the cohort design, it was vital to include all consecutive patients who met the inclusion criteria. The material studied was file-based, i.e. the risk assessments done for clinical purposes. Strict anonymity was insured in the handling of these data. Study V did not include special consent forms. A letter of information about the study accompanied the material sent out to the subjects. There were no patients included in the study and the act of sending back the questionnaires was considered as satisfactory consent.

The potential harm to the studied subjects should also be weighted against the usefulness of new knowledge that might result from the research. In this cost-benefit analysis, the main aim of the studies was to improve on assessments, management and decision concerning future risk of violence. Improved risk assessment procedures could strengthen the protection the protection of society and of the offenders, through more transparent procedures and better informed legal and clinical decisions, which would benefit potential victims, offenders and society at large.

To protect the integrity of subjects, the analysis and results were not presented in such a way that participants could be identified. Personal identity numbers were removed from all material as soon as research identity numbers had been given. In Study III, the expert assessors were not asked to rate their feelings towards their patients in the way that the nurse assessors were. Because of the small N (8), and the low number of experts (5). It was argued that persons, with knowledge about the conditions at the forensic centre, could deduct this sensitive information.

The materials of Studies I, III and IV were kept in secure conditions at the department of Behavioural Sciences at Linköping University. They were moved to secure conditions at The Forensic Psychiatric Centre in Vadstena in 2001, where materials from Studies II and V were kept from the start.

Studies I-IV all involved forensic patients. The designs of these were reviewed and approved by the Regional Research Ethics committee in Stockholm. Study I (#04-206/1), study II (#04.259/1), studies III and IV (#04-258/1). The committee commented about the lack of written informed consent (above) and that the reasons for this should be clearly stated. Since 2003, there is a written informed consent form used in the ongoing risk research program in Vadstena and persons, who do not consent will be excluded in future studies.



## RESULTS

Study I explored the inter-rater coherence and predictive accuracy of the HCR-20, version 1 (Webster et al, 1995). Because of the small N and the few points of observations (36) the range of correlations (Pearson's product-moment) between raters was calculated, rather than the mean correlation. The correlations were .76- .96 for the full 20 items, .85- .99 for the 10 historical items, .52- .95 for the 5 clinical items and .69- .94 for the 5 risk management items. Only the full scale and historical items were high enough for statistical significance ( $p < .01$ ). The content validity was estimated by comparing a crime index with the HCR-20. However item one of the HCR-20 was excluded, since it also assesses previous violent criminality. The correlation between the crime index and "HCR-19" was .67. The Crime Index reflected previous violence rather than outcome on future violent crime, the association between "HCR-19" and Crime Index was triangulated with previous research where the association between past and future violence was found to be .38 (Teplin et al, 1993).

In Study II, the base-rate of violent behaviour was high, overall 57% of the sample. The frequency of violent incidents was highest in the High Risk Management condition 0.12/week, lower in the Medium Risk 0.06/week and lowest in the Low Risk Condition, 0.03/week. The nature of violence also changed in the management conditions. There were increasing partial correlations for verbal and physical violence from high to low risk management conditions (.34, .47, and .82 respectively, controlling for the number of weeks in each condition). Violence to objects was associated with verbal and physical violence only in the low risk management condition, (.79 and .64). The overall prediction of all incidents of the HCR-20 was AUC .68, considered moderate predictive ability (Dolan & Doyle, 2000). The 10 historical items had the best predictive ability of the HCR-20 sub-scores, AUC .70. There was a tendency for the predictive ability of the HCR-20 and its subscales to increase as the risk management dropped. In the High risk management  $r = .21$ , AUC = .64, in Medium management  $r = .41$ , AUC = .83, and finally in Low management  $r = .50$  and AUC = .71. The HCR-20, total scale and the R-5 subset had predictive accuracy for violent reconvictions (AUC .84), as had Psychopathy (AUC .71). None of the risk subscales were associated with non-violent reconvictions. However the sample was small, only 16 of the 54 were reconvicted.

The next question was what might influence the structured risk assessment process. In Study III the nurses rated patients significantly higher than the experts (two-tailed T-test). The mean HCR-20 score for nurses was 26.3 (SD 6.10) while the experts' mean was 22.7 (SD 6.12). The same tendency was true for PCL: SV, the nurses mean rating was 16.1 (SD 6.1) and the experts' mean was 14.9 (SD 3.5). There was some coherence between nurses and clinicians, the overall correlation (Pearson's  $r$ ) was .59. The risk management items correlated at .36.

The nurses' HCR-20 scores were influenced by the scores on the feeling word checklist (FWL). The multiple regression, using HCR-20 scores as the dependent variable, was .659 and the R square was .433. ( $F = 4.79$ ,  $p = .005$ ). Whether the nurse was the appointed Key-nurse (staff-nurse) or not did not have a significant effect on

risk score (R square .10,  $F = 3.42$ ,  $p = .075$ ). The subscales on the feeling word checklist were associated with the risk scores. The feelings of “close” and “accepting” were associated with higher risk scores, while the feelings of “helpfulness” and “autonomous” were associated with lower scores ( $p < 0.05$ ).

The value attributed to different kinds of information, forming the basis of the HCR-20 assessment was also found to be significantly different between nurses and experts carrying out the assessments (Study IV). Nurses rated personal interaction as the most valuable source of information (44%); experts rated this as the least useful (14%). The experts rated the pre-sentence forensic evaluation report (FPE) as being the most valuable (39%); the nurses rated this as the second most valuable (26%). Experts rated the specific risk assessment interviews as the second most valuable source (31%), while nurses only attributed half this value to interviews (16%). Both nurses and experts rated hospital case-notes as the least useful source of information (15%, 16% respectively).

The outcome for the 8 patients in Study IV was followed for 48 months after discharge (conditional). 4 of them had recorded incidents of inpatient violence before release. 4 of the 8 were rearrested for a violent crime, 3 for common assaults and 1 for arson. A 5<sup>th</sup> person was rearrested for indecency. The nurses' HCR- scores were associated with inpatient incidents (Pearsons'  $r$ ). The full 20 score was associated at .63 ( $p < 0.1$ ) and the R-5 ratings at .76 ( $p < 0.05$ ). None of the nurses' assessment mean scores were associated with reconvictions.

The experts' HCR-scores were associated with reconvictions. Both the HCR-20 and the H-10 were associated at .81 ( $p < 0.05$ ) for violent reconviction. The experts H-10 was also associated with inpatient violence at .73 ( $p < 0.05$ ).

In paper V, the sex and age distribution of the participants was: 47% men amongst Clinicians, 53% men in the Criminal law professionals group, and 24% amongst Controls, thus with a clear trend towards more women in the Control group ( $\chi^2 [3] = 5.19$ ,  $p = .075$ ). The age distribution varied between the professional groups, so that Controls ( $M = 52.23$  years,  $SD = 6.48$ ) were slightly older than both Clinicians ( $M = 43.33$  years,  $SD = 9.68$ ) and Criminal law professionals ( $M = 47.12$  years,  $SD = 10.31$ ); one-way ANOVA,  $F [2] = 6.65$ ,  $p = .002$ .

The Clinicians scored more positive attitudes towards mentally disordered offenders (ATMDO) than Criminal law professionals and Controls (low scores are negative attitudes). The Criminal law professionals group was the most negative ( $M = 87.78$ ,  $SD = 12.53$ ). The Controls were intermediate ( $M = 89.58$ ,  $SD = 14.91$ ) and the clinicians were the most positive ( $M = 95.90$ ,  $SD = 14.77$ ); one-way ANOVA  $F [2] = 2.91$ ,  $p = .060$ ). There was no difference in ATMDO score between men ( $M = 90.78$ ,  $SD = 16.01$ ) and women ( $M = 91.06$ ,  $SD = 12.64$ );  $t [85] = .09$ , n.s. The correlation between age and ATMDO score was  $r = -.02$ , n.s. The correlation between age ( $M = 90.78$ ,  $SD = 16.01$ ) and women ( $M = 91.06$ ,  $SD = 12.64$ );  $t [85] = .09$ , n.s. The correlation between age and ATMDO was similar in all professional groups and among both men and women (details not shown).

There were no differences in inclination to discharge between Clinicians, Criminal law professionals and Controls as reflected by the first outcome measure, that is, on Question 1 in the vignette survey (Q1: “How inclined would you be to discharge the patient (if you were a member of the tribunal)?”

The inclination to discharge was slightly correlated to the attitudes towards mentally disordered offenders,  $r = -.326$  ( $p < .001$ ). The negative correlation is explained by the negative direction of the ATMDO, higher scores reflect more negative attitudes. The manipulation of perceived gain also had a relatively minor effect on the discharge inclination in a positive direction, the perceived gain condition made all groups slightly more inclined to discharge when perceived loss was held constant. Clinicians were less likely to discharge when the perceived gain was low than Criminal law professionals and controls.

A substantial effect of perceived loss was found equally for all three groups, when perceived gain was held constant. The mean inclination dropped by almost one point in the six-graded scale for all three groups in the high loss condition. This could be compared to the rise of inclination of about 0.2 of a point for high gain.

The main effects of the two variables “perceived gain” and “perceived consequences” (loss) were next put into ANOVA models. This revealed that only consequence had a clinically interesting effect on the outcome measure.

On the outcome measure Q1; how inclined would you be to release the patient, there was a significant effect of perceived consequence,  $\eta^2 = .58$ . For the variable perceived gain there was a very modest association,  $\eta^2 = .17$ . The attitudes questionnaire likewise had a modest association  $r = .326$  ( $p < .001$ ). Group membership had no significant effect. The inclination to discharge was also slightly correlated to the attitudes towards mentally disordered offenders,

On the second outcome (Q6, i.e. “How long do you think that NN needs to remain in secure conditions?”) only perceived consequence (loss) had a statistically significant effect ( $\eta^2 = .71$ ).

There were no interaction effects between the loss and gain conditions on either of the two outcome measures



## **DISCUSSION**

### **LIMITATIONS**

The concordance between risk assessors in Study I were unfortunately expressed as a range of correlations. From a methodological point, this would have been captured better by Cohen's  $\kappa$  (Cohen, 1960) or intra-class correlations (Shrout & Fleiss, 1979). Correlations do not capture the magnitude of differences in ratings and is not a good description of agreement. A further limitation of this study was the retrospective triangulation with previous violence as an indication of validity. Prospective studies are usually superior to retrospective designs. However, prospective designs are time consuming and expensive to conduct.

Study II had a broad definition of violent behaviour. It included self-harm, which was found to make up a third of the incidents. This was in an effort to actually capture all violent behaviour occurring, but it makes it hard to compare the data with other studies that uses definitions of violent threats (Nicholls et al., 1999) or a more limited definition of physical violence, as in the MacArthur dataset (Monahan et al., 2000).

In Study III, the expert-raters were not asked to rate their feelings, with the FWL, in the same way as the staff-raters did. Consequently, there was no way of comparing the magnitude of influence of feelings on the risk assessments. However, this limitation was accepted for two reasons. Firstly, the experts' HCR-20 assessments were made some months previous to the nurses' assessments, and as part of the clinical routine of the hospital. It was not considered meaningful to ask the experts how they felt about the patients in retrospect. Secondly, only one expert assessed each patient, and the outcome of the FWL-rating by the expert about the patient would have been obvious to anyone with local knowledge, who read the results. This was considered unethical, since the relationship between professional and patient is a sensitive issue. Study III did not include any outcome data, i.e. violent recidivism. The validity of the nurses' and/or experts' HCR-20 assessments could consequentially not be validated in Study III.

The limited number of patients and the limited number of expert-raters in studies III and IV are also obvious limitations. Study IV consisted of multiple case-reports, and was not a traditional empirical study. Rather, my ambition was with this study was to explore components of the risk assessment process that, I think, need further study.

Only two variables were explored in study V. The probability of re-offending was kept constant in the study, and the effect of probability statements could not be evaluated. To manipulate a further variable would entail a 2 x 3 x 3 design and would require a much larger number of subjects. This would be a very interesting topic for further research, perhaps using a similar design as study V.

## COMMENTS TO RESULTS

In Study I, there were indications of decent inter-rater coherence.. At the time of this study (1997) there were no published data on the inter-rater reliability of the HCR-20 scheme in the literature. The author felt that some rough guidance on this was needed before the assessment could be clinically implemented, hence the correlation coefficients. With time and during the implementation, other reports on reliability became available (Belfrage, 1998; Douglas et al., 1999) and the idea of recalculating the findings seemed superfluous. The content validity was defined retrospectively in a triangulation with previous research (Teplin, et al 1993). The reasoning behind comparing the “HCR-19” with past violent criminality, as defined by the “crime index scale” was based on the notion that nothing predicts future behaviour like past behaviour. This is sometimes referred to as *Meehl’s Malignant Maxim* (Meehl, 1971). In the clinical context of searching for risk assessment procedures, that were valid, but also manageable and comprehensive enough to be adopted and evaluated, it was felt that the triangulation in Study I, provided some indication of content validity that was seen as promising and warranting an adoption of the HCR-20 scheme as the standard risk assessment procedure of the Forensic Unit in Vadstena.

In spite of this “good enough” view there were a number of points that needed further study. The assumption of clinical and situational factors adding incrementally to the historical data needed further exploration. An overlap between the three sub-scores indicated the need for more elaborate analysis, for which the modest data of Study I were not adequate. To my knowledge this remains unattended in the literature. A factor analysis of a larger sample of HCR-20 protocols would be able to show whether the assumption of factors past, present and future actually corresponds in a factor structure, apart from being a clinically reasonable and appealing approach to violence risk assessment.

Study II showed that the base-rate of institutional violence was high, more than half behaved violently during the follow-up period. This was in line with previous research on inpatient violence in forensic psychiatric settings (Zeiss et al., 1996; Nicholls et al., 1999).

The study design had two advantages, rare in the risk assessment literature. The first being the use of a prospective design, the second being the use of a multitude of monitoring techniques, to get a fair estimate of the outcome criterion. The limitation was the limited number of available subjects. Particularly the re-offending was low, because of a limited number of subjects and the limited time-at-risk in the community. Although both PCL: SV and HCR-20 were found to be significantly predictive of violent reconvictions, this needs to be handled cautiously, and the sample will continue to be followed in time, in order to generate more definite results in the near future.

There were significant interaction effects found between risk prediction and management. No statistical predictive power of the HCR-20, was found in the high-risk management condition, while the predictive accuracy was medium to good in the medium and low risk-management conditions were high. Overall this was interpreted as supportive of the hypothesis that good management not only is worth doing, but also

tends to invalidate the violence prediction. It is “our task as clinicians to prove ourselves wrong when we predict risk” (Hart 1998, p.123).

It could be argued that this would be “obvious” and that empirical study of the link between violence prediction and risk management is superfluous and breaking into open doors. However, there are to my knowledge no studies that demonstrate the link, but more importantly it suggests that the HCR-20 may not be the best method to assess the risk of imminent, in-patient violence as suggested by some researchers (Douglas et al., 1997; Kroner and Mills, 2001; Nicholls et al., 1999; Webster et al., 2000).

The next threat to assessment accuracy investigated in this thesis was the influence of irrational factors on the assessment process. Study III suggested that nursing staff were greatly influenced by their feelings towards their patients when assessing violence risk with structured clinical methods (HCR-20). Because of the lack of comparative data for the expert group it couldn't be concluded that irrational influences are a threat to assessment accuracy, however the magnitude of the influence of feelings on variance in risk scores suggests that this is possible and warrants further study.

The issue of gathering information and how the rater's valued different sources of suggested different strategies between experts and staff-raters. The staff-raters quite clearly valued personal interaction most while experts did value forensic reports the highest. This could be interpreted in the model of Sawyer (1966) as though the experts used a more actuarial approach to data gathering, while nurses relied on less structured and more impressionistic way of gathering data. Unfortunately there is no way of knowing if this difference in approach also resulted in the differences in accuracy in assessing risk for violence.

Four of the eight patients in Study IV were subsequently reconvicted for violent crime. The frequency of reoffending is not interesting because of the selected, small sample. However, it is interesting to note that only the experts' HCR-20 assessments were associated with violent reconvictions. The staff-raters' HCR-20 scores were associated only with in-patient violence.

Another perspective on risk assessments was explored in Study V. This perspective concerned decisions made as a consequence of risk assessment. These decisions are actually about the further management of risk. How should the offender be handled? Should he be incapacitated, i.e. kept in secure conditions? Interestingly there were no differences in decision-making between criminal law professionals, clinicians or controls. Furthermore, attitudes to mentally disordered offenders made a small contribution to the inclination to discharge mentally disordered offenders. Instead a very large impact was found for the potential perceived loss from making a decision about discharge. The perceived gain by such a decision was substantially smaller in effect. In other words, the risk of making a false negative decision outweighs the gain perceived by making a true negative decision.

## DISCUSSION PROPER

The HCR-20 scheme may be criticised for the selection of risk factors to which it directs the clinicians' attention. It is obvious that, given a format of 20 items, there will have to be priorities made. However there are some factors that could be argued for inclusion from the general literature. Among static factors are parents' criminal history (Farrington, 1992), low IQ and neuropsychological dysfunction (Moffit, 1990). Among dynamic factors are criminogenic needs (McGuire, 2000) and others. My point here is that one should be aware of the relative arbitrary nature of the risk factors selected.

A common objection to risk assessment procedures is that it draws inference from groups of offenders and applies it to single individual cases (Grubin & Wingate, 1996; Söderström & Forsman, 2000). This problem with nomothetic data applied in an idiographic way presents a valid point. Clinicians are typically asked to do risk assessments of individual cases, not of groups of patients. However, to suggest that clinicians should refrain from making risk assessments on this ground is both naïve and counterproductive. Taken to its extreme, it would render most scientific research impotent and would fit badly with a science-practitioner paradigm (Douglas et al., 1999). The point is that information about statistical probabilities could and should be used in single cases (Dawes et al., 1989; Groove & Meehl, 1996). Meehl (1973) stated "the vulgar error is the cliché that 'we aren't dealing with groups, we are dealing with this individual case'" (p. 234). Einhorn (1986) argued that it might be a risky business to abandon decision rules on grounds that one might do better without them in an amply named paper "Accepting Error to Make Less Error". Ziskin and Faust (1988) warned for the same mistakes by emphasising the uniqueness of a case: "...the clinician is exploring territory where no one has explored before and for which no maps (prior knowledge) are available to offer guidance. Even so, there is an apparent tendency to believe that these unique features are extremely informative, or that they provide the basis of accurate prediction" (p 310). It is hard to think what the alternative to this position would be in the field of risk assessment, because all cases are unique, never rely on the odds, and never discharge forensic patients or discharge all patients?

However, even Paul Meehl agreed that there is a point in the individual uniqueness of patients and the nomothetic application on idiographic cases. Meehl (1973) named this the issue of the "broken leg countervailing". This refers to special cases where it could be wise to ignore the actuary of actuarial assessments. If one has an equation that would predict with a probability of .90 that a particular person would attend the cinema on Friday evening, one would be foolish not to accept the actuary on almost any occasion. However, if one receives information that the person has broken his leg on Friday afternoon, the applicability of the prediction rule would substantially reduce. Davis and Ogloff (2004) argued that clinicians have very little knowledge about when to override their actuarial assessments and replace them with clinical judgement or just plain common sense. They argued that this might contribute to the error variance of risk assessments and that the predictive accuracy may improve if there was such guidance. This approach warrants more examination and research. Webster and Eaves (1995) pointed out, in the original HCR-20 manual, that some cases should not be assessed. However, they refer more to the expertise of the assessor than to exclusion characteristics if a case.

The next question of some interest is the process of risk assessment. In most jurisdictions this task needs to be looked at in, at least, three sections: How, Who and What for?

The first question is how a risk analysis should be performed. Despite the wealth of scientific research into the area of risk assessment of mentally disordered offenders during the last decade, there is no generally accepted golden standard for how it should be done. (Monahan & Steadman, 1994; Monahan et al., 2001; Quinsey et al., 1998; Webster & Jackson, 1997; Webster et al., 1994). It is not only the choice of risk factors and the way to combine these that makes risk assessment a difficult task. There is also a lack of clarity about the acceptable professional standards that would guarantee the impartiality, quality and usefulness of the process of risk assessment (Heilbrun, 2004; Litwack, 1993; Litwack & Schlesinger, 1999). There are some suggested guidelines to assure the quality of risk assessments (Bonta, 2001; Douglas et al, 2001; Webster and Eaves, 1995). However there are issues regarding the process of risk assessment that are still understudied. Litwack (2002) highlighted the lack of adequate descriptive studies in areas of research in forensic psychology and psychiatry. Furthermore, Elbogen and colleagues (2002) called for *descriptive* scientific studies on how clinicians and other assessors perform their task. The authors contrasted this to the *prescriptive* research of the last decade, designed to evaluate the predictive validity of risk factors and risk assessment tools.

The second, “who” question has, in turn, several facets. What basic training is most suitable? Which professions should and should not perform assessments? What specific training in risk assessment should be required, if any? Is it appropriate that a person with a non-clinical background (e.g. experimental psychologists, criminologists) conduct risk assessments; should it be ring-fenced only to licensed mental health practitioners? What about the confusion of roles? Is it acceptable that mental health professionals assess clients with whom they also have a therapeutic responsibility (Litwack & Schlesinger, 1999)?

Study III and IV show that structured clinical assessment is context-specific. The nurse-assessors were greatly influenced by their feelings for the patient and also valued variable and unstructured information gathering. Interestingly, this resulted in HCR-20 scores valid for inpatient violence, but not for long-term violent recidivism. This could be perceived as a problem with the sheer amount of information considered in the assessment. The theory that structured risk assessments helps the assessor to structure the relevant cues for violence prediction seems to prevail (i.e. stick to the manual). They also favoured personal interaction as a source of information. The results of Studies III and IV suggest that biases and irrational influences need to be monitored and, if possible, controlled for, even in structured clinical risk assessments. Clinicians may be at risk of suffering from the same influences that makes “regular” clinical judgement invalid, “clinical judgement revisited” (Ziskin & Faust, 1988; Meehl, 1973).

The third question, “for what” concerns the decisions made on the basis of a risk analysis. In study V, it seemed that other factors than the probability of risk has an impact on these decisions.

In the risk research to little emphasis have been placed on protective factors, i.e. factors that might reduce risk, if present. Without some guidance to risk management strategies, the usefulness of risk guides ends in the synthesis or formulation. However, it would be a mistake to conceptualise risk assessment and management as two different entities or as different processes (Heilbrun, 1997). There is probably a reciprocal relationship, as suggested in Study II. In reality, we know little about what empirically constitutes protective factors, apart from stating that they are the mere absence of risk factors. It is almost a semantic question if the absence of certain risk factors should be reversed and hence viewed as protective factors. To ascertain data on what might constitute protective factors, studies needs to be designed to measure hypothetic protective factors as well as risk factors. In an innovative study, Haggård and colleagues (2001), interviewed “false positive” discharged forensic patients (i.e. high-risk individuals who did not recidivated violently).

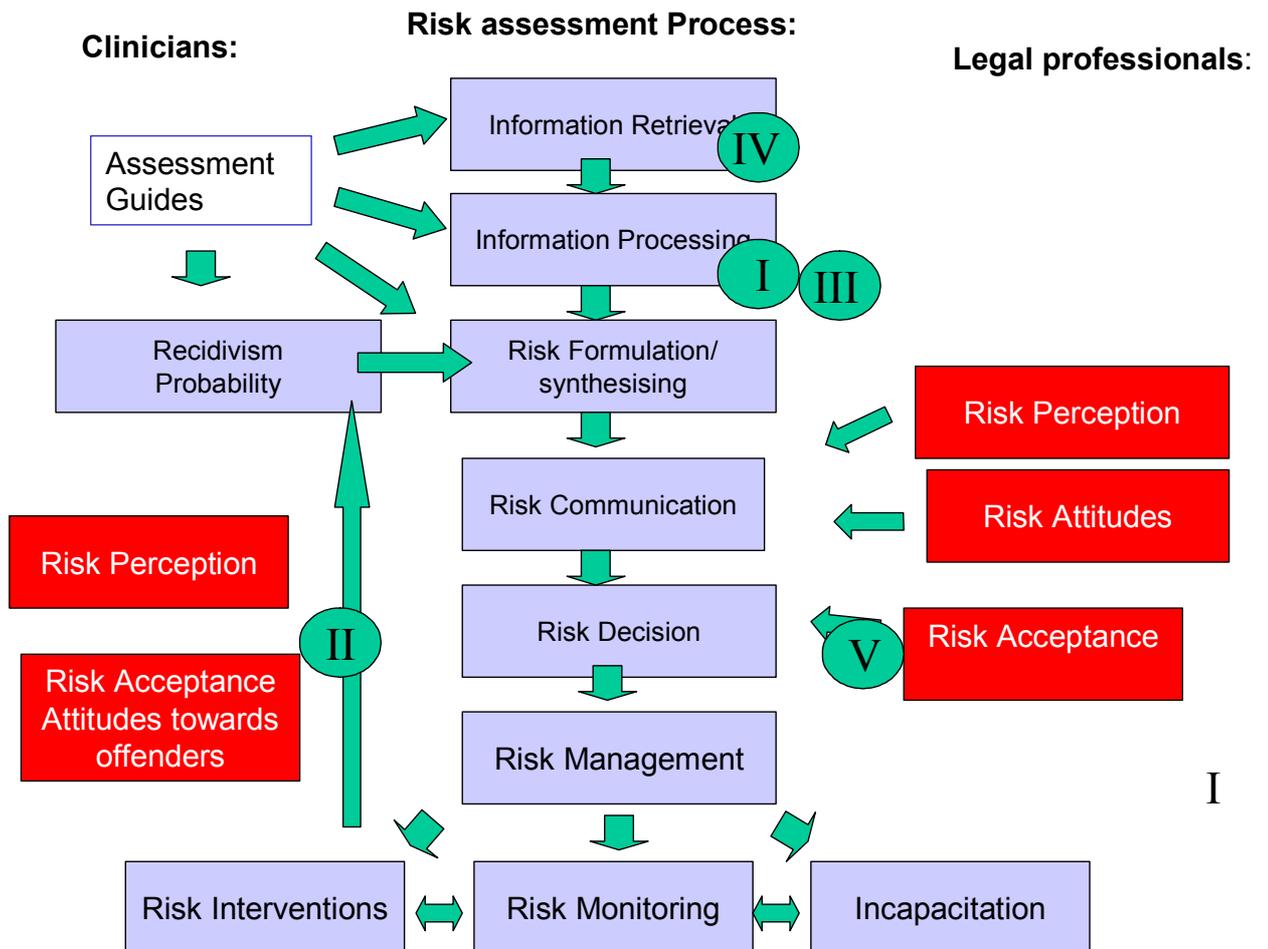
One model that include the assessment of protective factors as well as risk factors is the SORM (Structured Outcome Assessment and Community Risk Monitoring; Grann et al., in press). This is a tool for risk management rather than risk assessment, as reflected by the item setup and rating procedure. The factors included in the SORM are all dynamic (changeable), half of them contextual-dynamic and the other half individual-dynamic. Each factor is assessed in an actuarial manner for the absence of presence of the factors. In addition, each factor is clinically judged as to whether the presence of absence of the factors currently increases, decreases or does not affect risk (so-called risk effect assessment). Thus, in the SORM any factors can act as either as risk factor, protective factor or risk-irrelevant factors, depending on the unique circumstances that apply to the individual case. In an ongoing research project in Sweden, the SORM is currently used in released forensic patients who are monitored monthly with the SORM in the community. This approach might thus yield information on protective factors as well as risk factors.

Study II indicated that there is a loop back from risk management influencing the predictive ability of risk factors from risk guides. In this case, high-risk management seemed to invalidate predictions of violent risk. For future research, risk management strategies probably need more attention (Heilbrun, 1997). There is a risk for the focus of risk management is put on the incapacitation of the offender (as in discharge/not discharge decisions. Douglas and colleagues suggested, in the companion guide to the HCR-20 (2001), a conceptual model for risk management. This model suggests: monitoring, treatment, supervision and victim safety planning as the main strategies.

## **CONCLUSIONS**

A process approach to risk assessment and management could be useful in conceptualising the research in this thesis and in thinking about further research in the field. The five areas identified by Heilbrun (2004) could be expanded on and put into a process chart (Figure 1)

**Figure 1.** The risk assessment process.



Comment to Figure 1: Roman numbers refer to the area of interest in Stuyd I-V of the present thesis.

Guides like the HCR-20 are probably helpful in the first three facets of the process (Figure 1). They give guidance in the Information retrieval phase, by stating what information, regarding the patient, should be gathered and considered. In doing so, assessment tools might also inform us about what information could safely be left unattended, even if it would catch the attention of most clinicians, such as reports of depression, childhood sexual abuse and other psychologically interesting information, that would be of little use for a risk assessment.

However, the guides do not tell us about “the broken legs” (Meehl 1973), i.e. what are the factors that would inform us to disregard the actuary about the patient, and abandon the conventional risk assessment. A patient that tells you that he hates his wife, and will try to harm her when he gets the chance, should maybe not be told to hold on to that information, while we collect information about the age of his first violent episode or the degree of psychopathy. Recently Douglas and Ogloff (2003) showed that clinicians are more accurate in violence prediction, when they feel confident about their assessment (using the HCR-20). No risk assessment protocols, however, takes this into account. There may be a way to tap into the “broken leg” reasoning and increase the accuracy. A box on the protocol for the clinician, stating: “I’ve done the assessment by the book, but I don’t feel that this guy fits the form” might be helpful. Furthermore, guidelines and protocols do not tell us exactly from which sources to retrieve information. Different value could be attributed to various sources by the clinician (Study IV). A degree of variable data gathering (Sawyer 1966) is probably unavoidable in risk assessment.

Guides are also useful in the information-processing phase. They give guidance on how to combine the information gathered. Even if not presented as algorithms (although some actuarial risk appraisals are) they will give enough guidance to avoid impressionistic clinical judgement, so thoroughly dishonoured in the risk assessment literature Ennis & Litwack 1974; Monahan 1981). Fifty years ago Paul Meehl stated that we will “mainly use our heads, because there just aren’t any formulas. However, there is a strong argument to use formulas when they are available (Ziskin & Faust 1988). Even when adhering to a manual or guide, there are threats to the rationality of information processes. These are the biases that influence human decision-making and irrational influences (Study III). For the “who should do risk assessments” question these results indicates that professionals, who are trained in self- monitoring and observations might be the best suited for the job for this reason. In my opinion it should be clinicians, i.e. clinical psychologists and psychiatrists who conduct or supervise on risk assessments. They are after all, equipped to meet and speak to patients, as well as having scientific training.

Risk assessment guides have different ways of synthesising and formulating the risk assessments. Some focus merely on the probabilities of re-offending (Quinsey et al., 1998, Thornton & Hanson, 1999). Others, like the HCR-20 and the LSI-R tries to identify risk factors that are dynamic and could be targeted for intervention or monitoring. The HCR-20 also has room for the clinician to override the figures of the risk assessment in the overall level of risk statement (high, medium, low).

The communication and perception of risk are also influenced by other factors than the formulation of risk, whether in the form of odds ratio or relative risk. Attitudes to offender and particularly mentally disordered offenders could influence the risk perception. However, no effects of attitudes on risk acceptance were found in Study V. Violence risk decisions is an almost virgin field for scientific exploration (Monahan & Silver, 2003).

To conclude, I would like to suggest the following areas for further study:

- i. Definitions or framework of risk management that lends itself to comparative studies
- ii. The reciprocal relationship between risk assessment and management
- iii. The nature and validity of protective factors.
- iv. Descriptive explorations and recommendations for “golden standards” in risk assessment.
- v. Does training of the assessors have any effect for the accuracy of risk appraisal?
- vi. Ecological comparisons of processes of forensic decision-makers, factors that influence decisions of clinicians and judges.
- vii. Effects on decision-making of probability versus perceived loss.
- viii. Systematic ways of accounting for “broken leg” scenarios.



## **ACKNOWLEDGEMENTS**

It is not always easy for those around when middle-aged men get ideas about doing projects that they think is a good idea. Most of the acknowledgement goes to my family, Maria, Nils, Olle, and Axel for sticking by me in the process, despite all the travel, meetings and late nights.

I would also like to express gratitude to my supervisor and friend Martin Grann, who made this happen.

Many thanks to the “gang” i.e. the risk group in Vadstena for the many good times and laughs in the middle of a tough job, Ann-Marie Averland, Mikael Falkheim, Christer Jansson, Jan Lövestedt, Eva Petterson, Kenneth Rydenlund and others that participated through the years.

I would also like to thank the people at the Forensic Psychiatric Center in Vadstena. Kjell Andersson and Christina Karlsson for all support and for believing that R&D matters. Needless to say for the good sense of humour to the amigos at the R&D unit, Christer and Kenneth.

Also thanks to the people at the Centre for violence Prevention in Stockholm and especially to Niklas Långström and Anders Tengström, together with Martin Grann for making me think that risk research could be a good idea in a hotel room in Vancouver in 1999.

I would also like to thank colleagues at the department of Behavioural Science at Linköping University, special thanks to my co-authors Rolf Holmqvist and my ex-supervisor Rolf Sandell for your insight into methodology of measuring feelings.

I gratefully acknowledge the financial support for the risk project of the Forensic Psychiatric Centre in Vadstena, the dept. of Psychiatry at Linköping University Hospital, the Research council in Southeast Sweden, FORSS, and to the National board of Forensic medicine.



## REFERENCES

- American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, 47, 1597-1611.
- Andrews, D.A. (1995). The psychology of criminal conduct and effective treatment. In J. McGuire (Ed.), *What Works: Reducing Re-offending: Guidelines from research and practice*. Chichester: Wiley.
- Andrews, D.A., & Bonta, J.(1995). *The LSI-R: The level of Service Inventory- Revised*. Toronto: Multi-Health Systems.
- Andrews, D.A., & Bonta, J. (1998). *The Psychology of Criminal Conduct* (2<sup>nd</sup> edition). Cincinnati, OH: Anderson.
- Appelbaum, P. S. (1997). Ethics in evolution: The incompatibility of clinical and forensic functions. *American Journal of Psychiatry*, 154, 445-446.
- Augimeri, L., Koegl, C., Webster, C.D., & Levene, K. (1998). *The early assessment of risk list for boys (EARL-20B)* (Version 2) Toronto: Earls court Child and Family Centre
- Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Barrat, E. S., (1994). Impulsiveness and aggression. In J. Monahan, & H. Steadman (Eds.), *Violence and Mental Disorder: Developments in Risk Assessment* (pp 61-79). Chicago: University of Chicago Press.
- Baxter, R. (1997). Violence in schizophrenia and the syndrome of disorganisation. *Criminal Behaviour and Mental Health*, 7, 131-139.
- 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., 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., & 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 Services*, 1, 25-36.
- Berkowitz, L. (1964). Aggressive cues in aggressive behaviour and hostility catharsis. *Psychological Review*, 71, 104-122.

- Bilsky, W., & Winkel, F.W. (2002). *Fear of Crime, Personal Safety and Well-being: A common Frame of Reference*. Keynote paper at the 12<sup>th</sup> European Conference on Psychology and Law, European Association of Psychology and Law, Leuven, Belgium, September 14-17.
- Björkly, S. (1995). Prediction of aggression in psychiatric patients: A review of prospective prediction studies. *Clinical Psychology Review, 15*, 475-502.
- Björkly, S & Havik, O., (2003). TCO symptoms as markers of violence in a sample of severely violent psychiatric patients. *International Journal of Forensic Mental Health, 2*, 87-97.
- Blackburn, R. (1993). *The Psychology of Criminal Conduct*. Chichester, Wiley.
- Blackburn, R. (2000). Risk assessment and prediction. In: J. McGuire, T. Mason, & A. O’Kane (Eds.), *Behaviour, Crime and Legal Processes: A guide for forensic practitioners*. Chichester: Wiley.
- Boer, D.P., Hart, S. D., Kropp, P. R., & Webster, C.D. (1998). *Manual for the Sexual Violence Risk-20. Professional guidelines for assessing risk of sexual violence*. Vancouver Canada: The British Columbia Institute on Family Violence.
- Bonta, J. (2002). Offender Risk Assessment: Guidelines for selection and use. *Criminal Justice and Behavior, 29*, 355-379.
- Bonta, J., Law, M., & Hanson, R. K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin, 123*, 123-142.
- Boothby, J., & Clements, C. (2000). A national survey of correctional psychologists. *Criminal Justice and Behavior, 27*, 715-731.
- Borum, R. (1996). Improving clinical practice of violence risk assessment: Technology, guidelines and training. *American Psychologist, 51*, 945-956.
- Bowers, L., Nijman, H., Palmiserna, T., & Crowhurst, N. (2002). Issues in the measurement of violent incidents and the introduction of a new scale: the ‘attacks’ (attempted and actual assault scale). *Acta Psychiatrica Scandinavica, supplementum, 106*, 106-109.
- Brennan, P. A., Mednick, S. A., & Hodgins, S. (2000). Major mental disorders and criminal violence in a Danish birth cohort. *Archives of General Psychiatry, 57(5)*, 494-500.
- BRÅ, (2002). *Kriminalstatistik 2001*. Stockholm, Fritzes.

- Buchanan, A., & Leese, M. (2001). Detention of people with dangerous severe personality disorders: A systematic review. *The Lancet*, 358, 1955-1958.
- Centre for Disease Control and Prevention. (1994). *EpiInfo, Version 6* (computer software). Atlanta, USA: CDC Division of Surveillance and Epidemiology.
- Chapman, L. J. (1967). Illusory Correlation in observation report. *Journal of Verbal Learning and Verbal Behaviour*, 6, 193-217.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Cohen, L. (1983). The controversy about irrationality. *Behaviour and Brain Sciences*, 6, 317-370.
- Connolly, T., Arkes, H., & Hammond, K. (Eds.) (2000). *Judgement and decision making: An interdisciplinary reader* (2<sup>nd</sup> edn.) Cambridge: Cambridge University Press.
- Cooke, D. J., Michie, C., Hart, S. D., & Hare, R.D. (1999). Evaluating the screening version of the Hare psychopathy checklist-revised (PCL:SV): An item response theory analysis. *Psychological Assessment*, 11, 3-13.
- Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards an hierarchical model. *Psychological Assessment*, 13, 171-188
- Cullari, S. (1998). Treatment Resistance. In: S. Cullari (Ed.), *Foundations of Clinical Psychology*. (pp. 249-271). Needham Heights, MA: Allyn & Bacon.
- Davis, J. (Ed.) (1992). Group decisionmaking (special issue). *Organizational behavior and Human Decision Process*, 52.
- Davis, M., & Ogloff, J. R. P. (2004). *Broken leg countervailings and violence risk assessment: The views of Practicing Clinicians*. Paper presented at 4<sup>th</sup> annual conference of the International Association of Forensic Mental Health Services. Stockholm, June 6-9.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus Actuarial judgement. *Science*, 243, 1668-1674.
- Dawes, R. M. (1994). *House of Cards: Psychology and Psychotherapy built on myth*. New York: Free Press.
- Dernevik, M. (2002). Mentally Disordered Offenders in Sweden, In E. Blaauw, M. Hoeve, H. van Marle & L. Sheridan (Eds.), *Mentally Disordered Offenders; International Perspectives on Assessment and Treatment*, Hague, Elsevier.

- de Vet, H.C.W., van Tulder, M.W., & Bouter, L. M. (2003). Levels of evidence. Intellectual aid or absolute judgement? *Journal of Clinical Epidemiology*, *56*, 917-918.
- Dolan, M., & Doyle, M. (2000). Violence risk prediction: Clinical versus actuarial measures and the role of the Psychopathy Checklist. *British Journal of Psychiatry*, *177*, 303-311.
- Dolan, M., & Khawala, 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., 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., Cox, D. R., & Webster, C. D. (1999). Violence Risk Assessment: Science and Practice. *Legal and Criminological Psychology*, *4*, 149-184.
- Douglas, K. S., Ogloff, J. R. P., Nichols, T. L., & Grant, I. (1999). Assessing the 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., Hart, S. D., Eaves, D., & Ogloff, J. R.P. (2001). *HCR-20: Violence Risk Management Companion Guide*, Vancouver, B.C. Mental Health, Law, and Policy Institute, Simon Fraser University
- 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. (2003). The impact of confidence on the accuracy of structured professional and actuarial violence risk judgements in a sample of forensic psychiatric patients, *Law and Human Behavior*, *27*, 573-589.
- Einhorn, H. J. (1986). Accepting Error to make less Error. *Journal of Personality Assessment*, *50*, 387-395.
- Ennis, B. J & Litwack, T. R. (1974). Psychiatry and the presumption of expertise: Flipping coins in the courtroom. *California Law Review*, *62*, 693-752.
- Eronen, M., Hakola, P., & Tiihonen, J. (1996). Mental disorders and homicidal behavior in Finland. *Archives of General Psychiatry*, *53*, 497-501.

- Farrington, D. P. (1992). Explaining the beginning, progress, and ending of antisocial behaviour from birth to to adulthood. In J. McCord (Ed.), *Facts, Frameworks, and Forecasts: Advances in Criminological Theory, Volume 3*, New Brunswick: Transactional Publishers.
- Farrington, D. P. (1995). Psychosocial Influences on the Development of Antisocial Personality. In: G. Davies, S. Lloyd-Bostock, M. McMurrin, & C. Wilson (Eds.) *Psychology, Law, and Criminal Justice*. (pp 424-444). Berlin: De Gruiter.
- Farrington, D. P. (1996). The explanation and prevention of youthful offending. In J.D. Hawkins (Ed.), *Delinquency and crime: Current theories*. Cambridge: Cambridge University Press.
- Fazel, S., Danesh, J. (2002). Serious mental disorders in 23000 prisoners: A systematic review of 62 surveys. *The Lancet*, 359, 545-550.
- Fazel, S. & Grann, M., (2004a). The population impact of severe mental illness on violent crime. *Manuscript currently under review*.
- Fazel, S. & Grann, M., (2004b). Homicide offenders and psychiatric morbidity: A Swedish population study. *American Journal of Psychiatry*, 161,
- Folhälsoinstitutet (1995). *Ett folkhälsoperspektiv på våld. En beskrivning av våldet och det våldsförebyggande arbetet i Sverige. 1995*. Stockholm: Author.
- Friberg, I. (2003). *Enkät efter mordet på Anna Lindh*. Stockholm: Ingrid Friberg Samhällsinformation AB
- Gabriel, U. (2002). *Conceptual and Methodological Perspectives on the Psychology of Fear of Crime*. Paper presented at the 12<sup>th</sup> European Conference on Psychology and Law, European Association of Psychology and Law, Leuven, Belgium, September 14-17.
- Garb, H. (1992). The trained Psychologist as Expert Witness. *Clinical Psychology Review*, 12, 451-467.
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta analysis of the predictors of adult offender recidivism: What works! *Criminology*, 34, 575-607.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Gough, H. (1963). Clinical versus Statistical Prediction in Psychology. In L. Postman (Ed.), *Psychology in the making*. New York: Knopf.

- Grann, M., Långström, N., Tengström, A., & Kullgren, G. (1999). Psychopathy (PCL-R) predicts violent recidivism among criminal offenders with personality disorders in Sweden. *Law and Human Behavior, 23*, 203-215.
- Grann, M., & Nilstun, T. (2000). Rättspsykiatriska riskbedömningar är etiskt försvarbara (Forensic psychiatric risk assessments are ethically defensible). *Läkartidningen, 97*, 581-583.
- 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. (2002). *Riskbedömningar- Möjligheter och omöjligheter*. In: Statens Offentliga Utredningar, SOU 2002:26. Stockholm: Fritzes.
- Grann, M., & Fazel, S. (2004). Substance use and violent crime: Population study. *BMJ, 328*, 1233-1234.
- Grann, M., & Pallvik, A. (2002). An empirical investigation of written risk communication in forensic psychiatric evaluations. *Psychology, Crime & Law, 8*, 113-130.
- Grann, M., & Långström, N. (in press). Actuarial Assessment of Violence Risk: To weigh or not to weigh? *Criminal Justice and Behavior*.
- Grann, M., Sturidsson, K., Haggård-Grann, U., Hiscoke, U.L., Alm, P.O., Dernevik, M., et al. (in press). Methodological development: Structured outcome assessment and community risk monitoring (SORM). *International Journal of Law and Psychiatry*.
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (Mechanical, algorithmic) prediction procedures: The clinical- statistical controversy. *Psychology, Public Policy, and Law, 2*, 293-323.
- Grubin, D., & Wingate, S. (1996). Sexual offence recidivism: Prediction versus undersatnding. *Criminal Behaviour and Mental Health, 6*, 349-359.
- Gunn, J., Robertson, G., Dell, S. & Way, C. (1978). *Psychiatric aspects of Imprisonment*. London, Academic Press.
- Haggård, U., Gumpert, C., Grann, M. (2001). Against all odds: A qualitative follow-up study of high-risk offenders who were not reconvicted. *Journal of Interpersonal Violence, 16*, 1048-1065.
- Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristics (ROC) curve. *Radiology, 143*, 29-36.

- Hanson, R. K. (1998). What do we know about sex offender risk assessment? *Psychology, Public Policy, and Law*, 4, 50-72.
- Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348-362.
- Hanson, R. K., & Harris, A. J. R. (1998). *Dynamic predictors of sexual recidivism*. (User Report No. 1998 - 01). Ottawa: Department of the Solicitor General of Canada.
- Hanson, R. K. (2003). Personal communication, August 15<sup>th</sup> 2003.
- Hardman, D., & Harries, C. (2002). How rational are we? *The Psychologist*, 15, 76-79.
- Hare, R. D. (1991). *Manual for the Hare Psychopathy Checklist-Revised*. Toronto, Ontario: Multi Health systems.
- Hare, R. D. (1993). *Without conscience: the disturbing world of the psychopaths among us*. New York: Guilford Press.
- Hare, R. D., Clark, D., Grann, M., & Thornton, D. (2000). Psychopathy and the predictive validity of the PCL-R: an international perspective. *Behavioural Sciences and the Law*, 18, 623-645.
- Hare, R. D. (2003). *Manual for the Hare Psychopathy Checklist-Revised*. 2<sup>nd</sup> edition. Toronto, Ontario: Multi Health systems.
- Harries, C., & Harvey, N. (2000). Are absolute frequencies, relative frequencies, or both effective in reducing cognitive biases? *Journal of Behavioral Decision Making*, 13, 431-444.
- Harris, G. T., & Rice, M. E. (1997). Risk appraisal and management of violent behaviour. *Psychiatric Services*, 48, 1168-1176.
- Hart, S. D., Cox, D. N., & Hare, R.D. (1995). *Manual for the screening version of the Hare Psychopathy Checklist-revised (PCL:SV)*. Toronto, Canada: Multi Health Systems.
- Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and Criminological Psychology*, 3, 121-137.
- Hart, S.D., & Hemphill, J.F. (2002). Psychopathic personality disorder: Assessment and management. In E. Blaauw, & L. Sheridan (Eds.), *Psychopaths: current international perspectives*. Netherlands, The Hague: Elsevier.

- Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law and Human Behavior, 21*, 347-359.
- Heilbrun, K., Dvoskin, J., Hart, S. D., & McNiel, D. (1999). Violence risk communication: Implications for research, policy and practice. *Health, Risk & Society 1*, 91-106).
- Heilbrun, K. (in press) From risk assessment to risk management. In D. Carson & R. Bull (Eds.), *Handbook of Psychology in legal contexts* (2<sup>nd</sup> edition). Chichester: Wiley.
- Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology, 3*, 139-170.
- Hodgins, S. (1992). Mental disorder, intellectual deficiency, and crime. Evidence from a birth cohort. *Archives of General Psychiatry, 49*, 476 - 483.
- Hodgins, S., Mednick, S.A., Brennan, P.A., Schulsinger, F., & Engberg, M. (1996). Mental disorder and crime: Evidence from a Danish birth cohort. *Archives of General Psychiatry, 53*, 489-496.
- Holmqvist, R., & Armelius, B. (1996). The Patient's Contribution to the Therapist's Counter transference feelings. *Journal of Nervous and Mental Disease, 184*, 660-666.
- Holmqvist, R., & Armelius, B. (1996). The Patient's Contribution to the Therapist's Counter transference feelings. *Journal of Nervous and Mental Disease, 184*, 660-666.
- Holmqvist, R., & Armelius, B. (2000). Counter transference feelings and the psychiatric staff's self-image. *Journal of Clinical Psychology, 56*, 475-490.
- Holt, R. (1970). Yet another look at clinical and statistical prediction: Or is clinical Psychology worthwhile? *American Psychologist, 25*, 337-349.
- Howells, K. (1982). Mental Disorder and violent behaviour. In P. Feldman (ed.), *Developments in the study of criminal behaviour, volume 2*, Chichester: Wiley.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-291.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

- Kleinmuntz, K. (1984). The scientific study of clinical judgement in psychology and medicine. *Clinical Psychology Review*, 4, 582-591.
- Kraemer, H., Kazdin, A., Offord, D., Kessler, R., Jensen, P., & Kupfer, D. (1997). Coming to terms with the terms of risk. *Archives of General Psychiatry*, 54, 337-343.
- Kroner, D., & Mills, J. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*, 28, 471-489.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1995). *Manual for the Spousal Assault Risk Assessment Guide*. Vancouver, Canada: The British Columbia Institute on Family Violence.
- Leuger, R., & Petzel, T. (1979). Illusory correlation in clinical judgement: Effects of amount of information to be processed. *Journal of Consulting and Clinical Psychology*, 47, 1120-1121.
- Lindqvist, P., & Allebeck, P. (1990). Schizophrenia and Crime. A longitudinal follow-up of 644 schizophrenics in Stockholm. *British Journal of Psychiatry*, 157, 345-350.
- Link, B. G., & Stueve, A. (1994). Psychotic symptoms and the violent/illegal behaviour of mental patients compared to community controls. In J. Monahan, & H. J. Steadman (Eds.), *Violence and mental disorder* (pp 137-159). Chicago: University of Chicago Press.
- Litwack, T. R. (1993). On the ethics of dangerousness assessment. *Law and Human Behavior*, 17, 479-482.
- Litwack, T. R., & Schlesinger, L. (1999). Dangerousness risk assessments: Research, legal and clinical considerations. In A. Hess, & I. Weiner (Eds.), *Handbook of forensic psychology* (pp. 171-217). New York: Wiley.
- Litwack, T. R. (2001). Actuarial versus clinical assessment of dangerousness. *Psychology, Public Policy, and Law*, 7, 409-4443.
- Litwack, T. R. (2002). Some questions for the field of violence risk assessment: Or, "back to basics" revisited. *International Journal of Forensic Mental Health Services*, 1, 171-178.
- Main, D., Boon, J., & McAllister, H. (1996). Law and Psychology in the Perception of Crime Seriousness: Towards a Joint Approach. In G.

- Davies, S. Lloyd-Bostock, M. McMurrin, & C. Wilson (Eds.), *Psychology, Law and Criminal Justice: International Developments in Research and Practice*. Berlin: de Gruyter.
- McGuire, J. (2000) Explanations of criminal behaviour, In J. McGuire, T. Mason & A. O’Kane (Eds.), *Behaviour, Crime and Legal Processes*, Chichester, Wiley.
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.
- Meehl, P. E. (1957). When shall we use our heads instead of the formula? *Journal of Consulting Psychology*, 4, 268-273.
- Meehl, P. E (1971). Law and the fireside inductions: Some reflections of a clinical psychologist. *Journal of Social Issues*, 27.
- Meehl, P. E. (1973). *Psychodiagnosis: Selected papers*. Minneapolis: University of Minnesota Press.
- Melinder, K. (2002). A public health perspective on violence. *Acta Psychiatrica Scandinavica*, 106, 6-8
- Menzies, R. J., Webster, C. D., & Sepejak, D. S. (1985). Hitting the forensic sound barrier: prediction of dangerousness in a pretrial psychiatric clinic. In: C.D. Webster, M. H. Ben-Aron, & S. J. Hucker (Eds.), *Dangerousness: Probability and prediction, psychiatry and public policy*. Cambridge: Cambridge University Press
- Menzies, R. J., & Webster, C. D. (1989). Mental disorder and violent crime. In N. A. Weiner, & M. E. Wolfgang (Eds.), *Pathways to criminal violence*. Newbury Park, Ca: Sage.
- Milgram, S. (1963). Behavioural studies of obedience. *Journal of Abnormal and Social Psychology*, 67, 371-378.
- Moffit, T. E. (1990). The neuropsychology of juvenile delinquency: A critical review. In M. Tonry, & N. Morris (Eds.), *Crime and Justice*, 12, Chicago: University of Chicago Press, 99-169.
- Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. Beverly Hills, CA: Sage.
- Monahan, J. (1984). The prediction of violent behaviour: Towards a second generation of theory and policy. *American Journal of Psychiatry*, 141, 10-15.

- Monahan, J. (1988). Risk assessment of violence among the mentally disordered: Generating useful knowledge. *International Journal of Law and Psychiatry*, 11, 249-257.
- Monahan, J., & Steadman, H. J. (1994). (Eds.), *Violence and mental disorder: Developments in risk assessment*. Chicago, USA: University of Chicago Press.
- Monahan, J., Steadman, H. J., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., Silver, E., Roth, L. H., & Grisso, T. (2000). Developing a clinically useful tool for assessing violence risk. *British Journal of Psychiatry*, 176, 312-319.
- Monahan, J., Steadman, H. J., Silver, E., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., Roth, L. H., Grisso, T., & Banks, S. (2001). *Rethinking risk assessment: The MacArthur study of mental disorders and violence*. Oxford: Oxford University Press.
- Monahan, J., Heilbrun, K., Silver, E., Nabors, R., Bone, J., & Slovic, P. (2002). Communicating Violence Risk: Frequency Formats, Vivid Outcomes, and Forensic Settings. *International Journal of Forensic Mental Health*, 1, 121-126.
- Monahan, J., & Silver, E. (2003). Judicial Decision Thresholds for Violence Risk Management, *International Journal of Forensic Mental Health*, 2, 1-6
- Müller-Isberner, R., Jöckel, D., & Gonzales-Cabeza, S. (1998). *Die Vorhersage von Gewalttaten mit dem HCR-20*. Haina, Germany: Institut für Forensische Psychiatrie Haina
- Munkner, R., Haastrup, S., Joergensen, T., & Kramp, P. (2003). The temporal relationship between schizophrenia and crime. *Social Psychiatry Psychiatric Epidemiology*, 38, 347-353.
- Murphy, G., & Clare, I. (1995). Adult's capacity to make decisions affecting the person: Psychologists' contribution. In R. Bull, & D. Carson (Eds.), *Handbook of Psychology in Legal contexts*. Chichester: Wiley.
- Murray, C., & Lopez A. (1997). Global mortality. *The Lancet*, 349, 1269-1276.
- Nicholls, T. L., Vincent, G., Whittemore, K., & Ogloff, J. R. P. (1999). *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, Canada, November 6-10.
- Nijman, H., & Palmstierna, T. (2002). Measuring aggression with staff observation aggression scale-revised. *Acta Psychiatrica Scandinavica, supplementum*, 106, 101-102.

- Nilsson, A. (2000). Rättspsykiatriska farlighetsbedömningar har potential att bli vår tids "steriliseringsfråga" (Forensic psychiatric assessments of dangerousness have the potential of becoming the "sterilisation issue" of our time). *Läkartidningen*, 97, 1307-1308.
- Nordström, A., & Kullgren, G. (2003). Do violent offenders with schizophrenia who attack family members differ from those with other victims? *International Journal of Forensic Mental Health*, 2, 195-200.
- Nordström, A., Dahlgren, L., & Kullgren, G. (2004). *Victim relations and factors triggering homicides committed by offenders with schizophrenia*. Submitted manuscript, Umeå, Sweden: Umeå University.
- Norman, G.R., & Streiner, D.L. (2000). *Biostatistics: The bare essentials*. Hamilton, U.S.A.: BC Decker Inc.
- Novaco, R. (1994). Anger as a risk factor for violence among the mentally disordered. In J. Monahan, & H. Steadman (Eds.), *Violence and Mental Disorder: Developments in Risk Assessment* (pp 21-59). Chicago: University of Chicago Press.
- Nussbaum, D. (2004). *Theoretical and Empirical Problems in Applying the VRAG to Release Decisions and What should an Alternative look like?* Paper presented at the International Association of Forensic Mental health Services 4<sup>th</sup> Annual Conference, Stockholm, Sweden, June 6-9.
- O'Kane, A., & Bentall, R. P. (2000). Psychosis and offending. In J. McGuire, T. Mason, & A. O'Kane (Eds.) *Behaviour, Crime and Legal Processes*, Chichester, Wiley.
- Palmer, E. J. (2001). Risk Assessment: Review of Psychometric Measures. In D. Farrington, C. Hollin, & M. McMurrin (Eds.), *Sex and Violence: The psychology of crime and risk assessment*, London: Routledge.
- Paris, J. (1998). A biopsychological model of psychopathy. In T. Millon, E. Simonsen, M. Birket-Smith, & R. D. Davis (Eds.), *Psychopathy: Antisocial, criminal, and violent behaviour* (pp 277-287). New York: Guilford Press.
- Phares, J., & Trull, T. (1997). *Clinical Psychology* (5<sup>th</sup> ed.) Pacific Grove, CA: Brooks/Cole, 263-290.
- Philipse, M., deRuiter, C., Hildebran, M., & Bouman, Y. (2000). *HCR-20 Beoordelen van het risico van gewalddagig gedrag, Nederlandse Vertaling*, Prof. Mr W. P. J. Pompestichting en Dr. Henri van der Hoeven Stichting: Nijmegen en Amsterdam.

- Prins, H. (1986). *Dangerous Behaviour, The Law and Mental Disorder*. London: Tavistock.
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20*, 315-335.
- Quinsey, V. L., Rice, M. E., Harris, G. T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence, 10*, 85-105.
- Quinsey, V. L., Harris, G. T., Rice, M.E., & Cormier, C. A. (1998). *Violent Offenders: Appraising and Managing Risk*. Washington, D.C., American Psychological association.
- Redondo, S., Luque, E., & Funes, J. (1996). Social Beliefs about Recidivism in Crime. In G. Davies, S. Lloyd-Bostock, M. McMurrin, & C. Wilson (Eds.), *Psychology, Law and Criminal Justice: International Developments in Research and Practice*. Berlin: de Gruyter.
- Rice, M. E., & Harris, G. T. (1995). Psychopathy, schizophrenia, alcohol abuse, and violent recidivism. *International Journal of Law and Psychiatry, 18*, 333-342.
- Rock, D., Bransford, J., & Maisto, S. (1987). The Study of Clinical Judgement: An Ecological Approach. *Clinical Psychology Review, 7*, 645-661.
- Rosenhan, D. L. (1973). On being sane in insane places. *Science, 179*, 252-258.
- Seligman, M. (1995). The effectiveness of psychotherapy: The consumer reports study. *American Psychologist, 50*, 965-974.
- Salekin, R. T., Rogers, R., & Sewell, K. W. (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.
- Sawyer, J. (1966). Measurement and Prediction, clinical and statistical, *Psychological Bulletin, 66*, 178-200.
- Shah, S. A. (1978). Dangerousness: A paradigm for exploring some issues in law and psychology. *American Psychologist, 33*, 224-238.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability, *Psychological Bulletin, 86*, 420-428.

- Skeem, J. L., & Mulvey, E. P. (2001). Psychopathy and community violence among civil psychiatric patients: Results from the MacArthur violence risk assessment study. *Journal of Consulting and Clinical Psychology, 69*, 358-374.
- Slovic, P., & Monahan, J. (1995). Danger and coercion: A study of risk perception and decision making in mental health law, *Law and Human Behavior, 19*, 49-65.
- Slovic, P. (2000). *The perception of risk*. London, Earthscan Publications Ltd.
- Slovic, P., Monahan, J., & MacGregor, D. (2000). Violence risk assessment and risk communication: the effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior, 24*, 271-296.
- Stattin, H., & Magnusson, D. (1991). Stability and change in criminal behaviour up to age 30. *British Journal of Criminology, 31*, 327-346.
- Steadman, H. J., & Cocozza, J. J. (1974). *Careers of the criminally insane: Excessive social control of deviance*. Lexington, MA; Lexington books.
- Steadman, H. J., Mulvey, E. P., Monahan, J., Robbins, P. C., Appelbaum, P. S., Grisso, T., Roth, J. H., Silver, E. (1998) Violence by people discharged from acute psychiatric inpatient facilities and by others in the same neighbourhood. *Archives of General Psychiatry, 55*, 393-401.
- Steadman, H. J., Silver, E., Monahan, J., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., Grisso, T., Roth, L. H., & Banks, S. (2000). A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior, 24*, 83-100.
- Stone, A. A. (1985). The new legal standards of dangerousness: Fair in theory, unfair in practice. In C. D. Webster, M. Ben-Aron, & S. Hucker (Eds.), *Dangerousness: Probability and prediction, psychiatry and public policy* (p. 13-24). New York: Cambridge University Press.
- Sturidsson, K., Haggård-Grann, U., Lotterberg, M., Dernevik, M., & Grann, M. (2004). Clinicians' perceptions of Which Factors Increase or Decrease the Risk of Violence among Forensic Out-patients, *International Journal of Forensic Mental Health, 3*, 23-36.
- Susser, M., & Susser, E. (1996). Choosing a future epidemiology: II. From black box to Chinese boxes and eco-epidemiology. *American Journal of Public Health, 86*, 674-677.
- Swanson, J. W. (1994). Mental disorder, substance abuse, and community violence: An epidemiological approach. In J. Monahan, & H. Steadman (Eds.): *Violence and mental disorder*, p 101-136, Chicago: The university of Chicago Press.

- Szasz, T. (1974). *The myth of mental illness: Foundations of a theory of personal conduct*, (Rev. ed.). New York: Harper & Row.
- Söderström, H., & Forsman, A. (2000). Ethical and Methodological issues in forensic psychiatric risk assessments. In *Risk assessment in forensic psychiatric evaluation*. Report National Board of Forensic Medicine 200:1, Stockholm: Fritzes.
- Taylor, P. (1986). Psychiatric disorders in London's life sentenced offenders. *British journal of criminology*, 26, 63-78.
- Taylor, P., & Hodgins, S. (1994). Violence and psychosis: critical timings. *Criminal Behaviour and Mental Health*, 4, 267-289.
- Teigen, K. (2001). When equal chances = good chances: Verbal probabilities and the equiprobability effect. *Organizational behavior and Human Decision Process*, 85, 77-108.
- Tengström, A., Grann, M., Långström, N., & Kullgren, G. (2000). Psychopathy (PCL-R) as a predictor of violent recidivism among criminal offenders with schizophrenia. *Law and Human Behavior*, 24, 45-58.
- Tengström, A. (2001). Long-term predictive validity of the historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry*, 55, 243-249.
- Tengström, A., & Hodgins, S. (2002). Criminal behaviour of forensic and general psychiatric patients with schizophrenia: are they different? *Acta Psychiatrica Scandinavica*, suppl, 412, 62-66.
- Teplin, L.A., McClelland, G. M., & Abram, K. M. (1993). The role of mental disorder and substance abuse in predicting violent crime among released offenders. In S. Hodgins (Ed.), *Mental Disorder and Crime*. Newbury Park, CA: Sage.
- Thornberry, T. P., & Jacoby, J. E. (1979). *The criminally insane: A community follow-up of mentally ill offenders*. Chicago, IL: University of Chicago Press.
- Thornton, D., & Hanson, R. K. (1999). *The Static 99: Assessing Risk for Sexual Reoffending*. London: The Home Office.
- Tiihonen, J., Isohanni, M., Räsänen, P., Koiranen, M., & Moring, J. (1997). Specific major mental disorders and criminality. A 26 years prospective study of the 1966 Northern Finland birth cohort. *American Journal of Psychiatry*, 154, 840-845.
- Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.

- Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the psychopathy checklist factor scores: A meta-analysis. *Law and Human Behavior*, 27, 541-558.
- van Tijen, N., & Verheul, R. (2002). The Origins of Psychopathy. (pp. 39-59) In E. Blaauw, & L. Sheridan (Eds.), *Psychopaths: Current international perspectives*. The Hague: Elsevier.
- Webster, C. D., & Ben-Aron, M. H. (1985). Dangerous doctor Fisher: a case history in clinical assessment? In C.D. Webster, M. H. Ben-Aron, & S. J. Hucker (Eds.), *Dangerousness: probability and prediction, psychiatry and public policy*. Cambridge: Cambridge University Press.
- Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C. A., & Quinsey, V. L. (1994). *The violence prediction scheme: Assessing dangerousness in high-risk men*. Toronto: University of Toronto, Centre of Criminology.
- Webster, C. D., & Eaves, D. (1995). *The HCR-20 Scheme: The assessment of Dangerousness and Risk*. Vancouver: Simon Fraser University.
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20. Assessing Risk for Violence. Version 2*. Vancouver: Mental Health, Law and Policy Institute.
- Webster, C. D., & Jackson, M. (Eds.), (1997). *Impulsivity: Theory, assessment, and treatment*. New York: Guilford.
- 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.
- Webster, C. D., & Bailes, G. (2001). Assessing violence risk in mentally and personality disordered individuals. In C. Hollin (Ed.), *Handbook of Offender Assessment and Treatment* (pp. 71-84). Chichester: Wiley.
- Webster, C. D., Müller-Isberner, R., & Fransson, G. (2002). Violence risk assessment: Using structured clinical guides professionally. *International Journal of Forensic Mental Health*, 1, 185-193.
- Weinberger, J. (1995). Common factors aren't so common. *Clinical Psychology: Science and Practice*, 2, 45-69.
- Whyte, C., Constantopoulos, C., & Bevans, H. (1982). Types of countertransference identified by Q-analysis. *British Journal of Medical Psychology*, 55, 187-201.

Von Hofer, H. (2000). Criminal violence and youth in Sweden: a long Yterm perspective. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 1, 56-72.

World Health Organisation (2002). *World report on violence and health*, Geneva: Author.

[www.sfu.ca/psychology/groups/faculty/hart/violink.htm](http://www.sfu.ca/psychology/groups/faculty/hart/violink.htm)

Yerushalmy, J. (1947). Statistical problems in assessing methods of medical diagnosis with special reference to X-ray techniques. *Public Health Reports*, 62, 1432-1449.

Zeiss, R., Tanke, E., Fenn, H., & Yesavage, J. (1996). Dangerousness commitments: Indices of future violent potential. *Bulletin of the American Academy of Psychiatry and Law*, 24, 247-253.

Ziskin, J., & Faust, D. (1988). *Coping with Psychiatric and Psychological Testimony*. (4<sup>th</sup> ed). Marina del Rey, CA: Law and Psychology Press.