PREDICTING VIOLENT RECIDIVISM OF TREATED VIOLENT OFFENDERS USING THE PSYCHOPATHY CHECKLIST-REVISED AND THE VIOLENCE RISK SCALE

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Arts

in the Department of Psychology

University of Saskatchewan

Saskatoon

By

Grant N. Burt

Spring 2000

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master of Arts degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department of Psychology or the Dean of the College of Graduate Studies and Research. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5A5

Abstract

The present study compared the ability to predict violent recidivism using the PCL-R (Psychopathy Checklist- Revised) and the VRS (Violence Risk Scale). The study examined post treatment violent convictions of 60 federal offenders who had participated in a high intensity violence reduction correctional treatment program. The prediction of presence or absence of violent recidivism, and the cumulative number and rate of violent convictions at 1, 2, 3, 4, and 5 years follow up was investigated. VRS ratings of change in risk after treatment were not found to provide a predictive improvement over VRS pre treatment ratings. Correlational, simple regression, and ROC (Receiver Operating Characteristics) analysis indicated that the PCL-R demonstrated a stable relationship to violent recidivism, while the VRS provided a stronger prediction of risk in the short term (i.e., 2-3 years follow up) but was generally unrelated to violent recidivism at a longer follow up period (i.e., 4-5 years follow up). It is suggested that these results reflect the static and dynamic theoretical approaches of the PCL-R and VRS, respectively. Implications of this study indicate that comprehensive file information may be necessary to assess changes in risk accurately. In addition, the differences in the predictive ability of the VRS over length of follow up suggests caution for comparison of static and dynamic risk measures in future research.

ACKNOWLEDGEMENTS

I want to acknowledge the many individuals who provided their assistance and unwavering support in my research. Foremost, I wish to express my gratitude to Dr. Stephen Wong, my supervisor, who provided the insight, encouragement and dedication to making the research process more fulfilling and rewarding for me. His intellectual guidance, his sense of humor, and his continued support and understanding, provided for an excellent working relationship and instilled in me a better appreciation for the complexity of the research process. He has facilitated my development and enhanced the many research skills I will undoubtedly find essential in future years. I wish to express my sincere thanks to him for providing a positive experience throughout the thesis process, and for presenting such a good model of how to provide quality of supervision to students.

I wish to extend my gratitude to the other members of my thesis committee for their helpful input which has greatly contributed to increasing the quality of thesis experience, in addition to the thesis itself. I also wish to express my thanks to the administrative and clerical staff at the Regional Psychiatric Centre for providing me with the opportunity to explore the present topic. In particular, I wish to thank Gina Toni and Sonia Dhaliwal for their assistance in data collection and VRS coding. I also wish to thank Kevin Rice, for his help collecting, organizing and coding data, in addition to providing feedback and support throughout the research process.

Finally, I wish to thank my family and friends for their support. My sincere thanks and deep gratitude to my Mom and Dad for the love and words of encouragement you have given me throughout all my endeavors and challenges.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Violent crime has a substantial financial and human cost to society. As the majority of violent crime is committed by a small portion of the criminal population (Greenberg, 1996), understanding, predicting, and reducing violent recidivism is a critical component in the overall violence reduction strategy. Therefore, a primary focus of psychologists, criminologists and recent public policy has been on the management and the prediction of violence among habitual or repeat violent offenders (Simon, 1996). Subsequently, one central role which forensic psychologists must undertake is to assess violence risk of potential or known violent offenders.

The already overburdened and "financially strapped" criminal justice system (Greenwood et al., 1996), increasingly is facing more public pressure to demonstrate less leniency and more punitiveness towards offenders (Douglas, Macfarlane & Webster, 1996; Monahan, 1996). Decisions on sentencing, statutory release, eligibility for parole, conditions for parole, and type of appropriate treatment all depend on evaluations of the offender's level of dangerousness. Assessments of dangerousness, which frequently rely on predictions made by mental health professionals, must be able to distinguish between those criminals who pose a future violent risk to the public, and those who do not. With increasing pressure for mental health professionals to predict violence with offender populations, psychological research and assessment techniques have evolved rapidly, with a central focus on those factors which demonstrate a strong empirical link to violence.

In section 1, the evolution of risk prediction will be presented based on a model discussed Monahan (1996). Section 2 summarizes and presents a comparison of two contemporary risk prediction instruments, the Psychopathy Checklist- Revised (Hare, 1991) and the Violence Risk Scale (Wong & Gordon, 1996) and the research hypotheses. Section 3 is the methodology of the research including subject selection criteria, measures of predictive validity, and the psychometric characteristics of the Violence Risk Scale. Section 4 is the results of the statistical comparison of the two risk instruments in predicting violent recidivism over several follow up periods. Section 5 discusses the implications of the results and suggests new directions for future research.

2. LITERATURE REVIEW

2.1 Risk Assessment and Prediction

2.1.1 Risk Assessment: First Generation

With more frequent requests for psychologists to assess the dangerousness of individuals, both inside and outside of correctional settings, there has been an increased need for precision in risk assessment. Historically, mental health professionals have relied on clinical judgement for risk assessment, however; it has become increasingly clear that this approach is grossly inadequate (Bonta, Law & Hanson, 1998; Kozol, Boucher & Garofalo, 1970; Monahan, 1996). In his review of the 1960s and 1970s literature, Monahan (1980) referred to risk assessment based on clinical judgement as "first-generation" research, which he demonstrated not only to be inadequate, but suggested that such practice may have led to unethical decision making. In a later review, Monahan (1996) indicated that clinical risk predictions were so inadequate that they were challenged as having the legal effect of violating the Fourteenth Amendment's due process or equal protection clauses, and/or the Eighth Amendment's prohibition of cruel and unusual punishment. Monahan (1980) explains that because of the very low base rate of violence, even among criminal populations, there is inevitably a large bias to over-predict violence (i.e., to predict individuals will behave violently when in actuality they do not). He indicated some clinicians' predictions of violence were only about 20% to 35% accurate, that is, 65% to 80% of individuals who were predicted to behave violently in the future demonstrated no future violence in a 4 to 5 year follow up. He

concluded that the validity of first generation predictions of violence (i.e., decisions based on professional judgements) is dubious at best.

2.1.2 Risk Assessment: Second Generation

With the growing concerns over the inaccuracy of clinical judgement in predicting violence, attention focussed on utilizing more objective means to assess risk. In particular, actuarial or static variables which were empirically related to general and violent recidivism became the basis of risk prediction. These variables are primarily historical or unchangeable features of an offender (e.g., age at first conviction). Static variables which demonstrated a strong empirical relationship to violent recidivism were incorporated into rating scales or assessment checklists in order to provide a basis for risk prediction. This process led to the development of empirically validated and reliable psychometric risk instruments which allowed mental health practitioners to greatly improve their predictions of violence (Otto, 1992). These instruments have been referred to as "second generation" risk instruments (Monahan, 1980). In general, second generation risk instruments have provided a host of empirically tested variables that are useful in identifying high risk offenders.

One of the most widely used second generation risk instruments has been the General Statistical Information on Recidivism Scale (GSIR) which was developed by Nuffield (1982). The predictive validity of the GSIR was based on a random sample of approximately 2500 male releasees from Canadian federal penitentiaries between 1970 and 1972 inclusive. The GSIR identified fifteen static variables which were closely related to general recidivism (e.g., current offense, longest prior time period without the commission of an offense, etc.). Subsequent modification of the weighting of particular GSIR items (Nuffield, 1982) has further improved the GSIR's predictive ability.

Although the GSIR and other second generation risk prediction instruments are still used today, and have shown a marked improvement over clinical judgement (Otto, 1992), they have three important limitations. First, second generation instruments generally lack a theoretical background, and subsequently, cannot generally be used to develop our understanding of recidivism, or to provide directions for improving risk prediction. Second, although second generation instruments have demonstrated a strong relationship to general recidivism, overall they have only a modest relationship to violent recidivism (Bonta, Harman, Hann & Cormier, 1996). Finally, these measures rely entirely on static, or historical variables, which are not susceptible to change. Static measures assess an offender's level of risk as a constant attribute, despite the recognition of the important influence of situational or dispositional factors. A more accurate and comprehensive approach to risk prediction must incorporate dispositional and situational influences into risk predictions, rather than relying solely on a static measurement of risk.

Using static predictors as the basis of risk assessment is highly problematic, particularly in settings in which dispositional changes or situational changes would be expected (e.g., when offenders have successfully completed intensive correctional treatment/therapy, or when they are released under high levels of supervision). The extensive use of static variables in second generation risk assessment may be due to three important factors.

First, data based on static variables are easy to gather in that the data are often readily available in case files, and require little, if any, subjective judgement or clinical skills.

Second, research using static variables has identified several strong predictors of future violence. For example, the Violence Risk Appraisal Guide (VRAG), which is comprised entirely of static predictors, has yielded significant correlations with violent recidivism in forensic research (Harris, Rice & Quinsey, 1993). Clearly, static factors have an important story to tell in risk prediction, and are not used simply for convenience.

Finally, there may be an underlying assumption of second generation instruments that many assessors find appealing. A reliance on a static approach to predicting violence suggests that changes in risk are unlikely or irrelevant, and can, to a large extent, be ignored by the assessor. This approach has serious implications for treatment programs, and is reminiscent of Martinson's (1974) conclusion that "nothing works". Even after Martinson (1979) recanted his initial assertion that correctional treatment was ineffective, and despite subsequent meta-analyses of research with adult offenders which has demonstrated strong treatment effects (Andrews & Wormith, 1989; Andrews et al., 1990), there is still considerable debate concerning the efficacy of treatment with offenders. Perhaps the assumption that "nothing works" with offenders has grudgingly changed to the conception that "nothing works" with violent offenders.

2.1.3 Risk Assessment: Third Generation

Third generation risk assessment instruments are theoretically based rather than empirically derived, and therefore provide a substantial conceptual advantage over second generation instruments. A theoretical framework provides a means to understand differences between offenders and, therefore, may be particularly important in differentiating between habitually violent offenders and other types of offenders. Assessments without a theoretical framework suggest a failure to recognize the

underlying causes for the heterogeneous nature of offender populations. This tendency is perhaps most apparent with the diagnosis of Antisocial Personality Disorder (APD) which has been claimed to apply to approximately 50% to 75% of male federal offenders (Hare, 1996). Such a high prevalence rate among offenders suggests that APD may be essentially synonymous with criminality and is thus non-discriminating as an explanation of violent criminal behavior.

As indicated earlier, a major shortcoming of second generation risk instruments is the absence of dynamic or changeable risk factors. The failure to recognize the importance of dynamic factors is especially distressing in light of a recent meta-anlaysis which demonstrates that dynamic factors were at least as effective as static factors in predicting general recidivism (Gendreau, Little & Goggin, 1996). Therefore, a second development in third generation risk measures is the conceptualization of risk as a dynamic construct.

2.2 Third Generation Risk Instruments

2.2.1 The Psychopathy Checklist-Revised (PCL-R)

One instrument which has empirically demonstrated success in identifying high-risk offenders, and which has a strong theoretical basis for understanding why some offenders may behave violently is the Psychopathy Checklist-Revised (Hare, 1991). The PCL-R is perhaps the most notable of the third generation risk instruments, particularly considering its strong influence on the development and direction of contemporary forensic research.

Psychopaths have been described in the psychological literature as "intraspecies predators who use charm, manipulation, intimidation, and violence to control others and satisfy their own needs. Lacking in conscience and in feelings for others they cold-

bloodedly take what they want and do as they please, violating social norms and expectations without the slightest sense of guilt or regret" (Hare, 1996, p.25). Although this modern description is largely based on Cleckley's (1976) original conception of psychopathy, Hare has argued that until recently, there has not been a uniform method of identifying, much less understanding, the criminal psychopath (Hare, 1996).

Based on the sixteen traits identified by Cleckley (1976), Hare designed the Psychopathy Checklist (PCL), a 22 item checklist to assess what he considered to be psychopathic traits (Hare, 1980). Ten years after the publication of the PCL, Hare modified the original checklist to include only 20 items (PCL-R), in an effort to reduce some of the redundancy found in the original PCL (Hare, 1996). The PCL and PCL-R consist of items which are rated based on a semi-structured interview and a review of detailed collateral or file information. In general, psychopathic traits are rated in the interview by observing the behavior and interaction style of the offender, rather than through the content of the interview (Hare, 1996).

Each PCL-R item is rated as either a 0, 1, or 2. A "zero" rating indicates the offender is unlike the characteristic described by the item, a "one" indicates that the offender is somewhat like the characteristic described by the item, and a "two" indicates that the offender closely resembles the characteristic described by that item. Therefore, the maximum possible PCL-R rating is 40. Hare (1996) states that scores above 30 are indicative of psychopathic offenders in forensic populations, although a rating of 25 can be used as a decision point for research purposes. Hare stresses that these guidelines are somewhat arbitrary, and should not rigidly be applied or interpreted. A copy of the PCL-R rating sheet is included in Appendix A.

Although the PCL and PCL-R were originally designed to assess psychopathy, conceptually, many of the psychopath's characteristics (i.e., impulsivity, lack of empathy, feelings of grandiosity, tendency to manipulate) are also indicative of an active criminal lifestyle and general disregard for authority. The relationship of the PCL / PCL-R and criminal recidivism has been quite impressive, perhaps indicating that, at least informally, the PCL should be considered a "bench mark" in risk prediction.

Several researchers have demonstrated that offenders with high scores on the PCL-R committed more crimes than the average criminal offender (Hare, 1981; Hare & Jutai, 1983; Wong, 1984), and that psychopathic offenders recidivated with a violent crime 3.5 times more often than did nonpsychopathic offenders (Hare & McPhearson, 1984). In addition, psychopaths have been reported to have a higher rate of institutional violence, and to be more likely to attribute hostile intent in hypothetical ambiguous situations than other violent offenders (Serin, 1991). In terms of conditional release, PCL scores also were able to predict success / failure (i.e., $r = .33$) of maintaining release status (Hart, Kropp & Hare, 1988). For conditional release, PCL prediction was superior to predictions based on relevant criminal history and demographic variables (Hart et. al., 1988). In addition, a much greater proportion (77%) of offenders who were assessed as psychopathic reoffended with another violent crime compared to offenders who were assessed as non-psychopathic offenders (22%), even when the groups were matched on past history of violence and number of prior violent offenses (Harris, Rice & Cormier, 1989). A comparative study (Harris et al., 1993), using the PCL-R and 12 other recognized predictors of violent recidivism, found that the PCL-R provided the best predictive value (i.e., $r = .34$) for violent recidivism. In general, the correlation

between the PCL-R and violent recidivism has been superior to comparable risk instruments.

In addition to the predictive advantage provided by the PCL-R, the construct of psychopathy also introduces a theoretical framework to understanding the habitual violent offender. Theoretically, individuals who possess the psychopathic personality traits (e.g., lack of empathy, grandiosity, manipulative nature) may be expected to be more likely to possess the respective behavioral characteristics (e.g., impulsivity, need for stimulation, parasitic lifestyle) and to engage in a more active criminal lifestyle. Consistent with this model, a well replicated oblique two-factor solution to the PCL-R has demonstrated that the personality and behavioural aspects of psychopathy are distinct, but related (Hare, 1991). The core personality items of the PCL-R (8 items) load strongly on the first factor, and the anti-social behavioral characteristics of the PCL-R (9 items) load strongly on the second factor. The strong correlation (i.e., $r = .50$) between Factor 1 and Factor 2 have led to the claim that it is the core "personality" features of the psychopath that make them more susceptible to engaging in criminal behavior (Hare, 1991).

Despite the demonstrated predictive validity of the PCL-R, and the theoretical background it provides to understanding violent recidivism, the assessment of psychopathy, much akin to second generation risk instruments, is based on the ratings of static and unchanging factors. Although the PCL-R brings a conceptual framework to understanding the persistent or habitual violent offender, it still relies on the degree to which an offender matches a relatively stable prototypical psychopathic personality (Hare, 1996). Because the PCL-R is based on the measurement of stable global personality traits and enduring antisocial behaviors to distinguish effectively between

low, medium and high risk offenders, it presents risk as stable and constant. In particular, considering that PCL-R ratings are based on an offender's life history, even dramatic changes in behavior during treatment would be unlikely to influence PCL-R ratings because of the relatively short time span of correctional treatment programs (i.e., 2 - 6 months).

The static approach used by the PCL-R is not surprising in that Hare (1996) has argued that the PCL-R was designed to assess a personality disorder, rather than to be used to predict violent recidivism. Although not developed specifically to assess criminal risk, the PCL-R is widely used in criminal justice settings to guide forensic assessments. The PCL-R has been utilized to assist in parole decisions, risk assessments, assessing statutory release times, and to determine suitability for correctional programming (Hare, 1991). The predictive utility of the PCL-R appears to lend itself well to most forensic settings, despite its static nature.

Although there has been a general move away from assessing historical and static variables in risk assessment, many professionals and researchers continue to conceptualize risk primarily in a static way, as evidenced by the popularity and extensive use of the PCL-R in forensic settings and research. This tendency may limit the development and accuracy of risk prediction measures and impede our understanding of the factors that may influence risk levels.

2.2.2 The Violence Risk Scale (VRS)

Originally developed as a research tool, the VRS has been designed to provide guidance in risk assessment with violent offenders who have completed correctional treatment. The VRS assesses 6 static and 20 dynamic risk factors which are highly related to violent recidivism, and has demonstrated good postdictive validity with

federal offender populations (Wong & Gordon, 1996). The VRS dynamic variables measure risk factors which are susceptible to change (e.g., criminal attitudes, interpersonal aggression, substance abuse). The conceptualization of variables as dynamic allows a rater to adjust the risk level of each variable if the rater determines that situational or dispositional changes have influenced an offender's likelihood to behave violently.

Although the VRS is the only dynamic risk assessment instrument designed specifically to assess treatment changes, the importance of dynamic variables in risk prediction has become increasingly evident. For example, another risk assessment instrument which can measure changes in risk is the Level of Service Inventory-Revised (LSI-R; Andrews, 1982). It comprises 54 dynamic and static items which are grouped into subcategories representing 10 risk areas. The HCR-20 (Webster & Eaves, 1993), designed for the assessment of risk in criminal and psychiatric populations, also utilizes both static and dynamic factors.

The VRS can be distinguished from other dynamic risk instruments primarily because of its strong emphasis on dynamic risk variables (i.e., 20 dynamic items, 6 static items) and because of its focus on assessing changes in risk with offenders who have completed treatment. For professionals working within a treatment setting, the VRS provides the necessary flexibility to assess changes in risk while still providing a reliable and theoretically meaningful basis to understand these changes (Toni, Wong & Burt, 1999). A copy of the VRS rating sheet is provided in Appendix B.

2.2.2.1 Pre treatment Assessment Using the VRS

The VRS is an evolving research tool. The current VRS pre treatment rating system (Part A) provides a comprehensive evaluation of violent risk based on static and

dynamic items. It should be noted that several of the dynamic risk items have been previously considered as static predictors in past research. For example, "martial status" has been traditionally identified as an important static risk item (Harris et al., 1993), with married offenders being at a lower risk for violent recidivism than their single counterparts. Although "martial status" is not a "true" static item, in that it is susceptible to some change, it is relatively stable in offender populations. The VRS uses a dynamic item, "Stability of Relationships with Significant Others" (item 13), to measure the underlying dynamic construct related to "martial status". Stability of relationships with significant others can change and is amenable to treatment. These changes may be indicative of better emotional control, higher levels of stability, and a greater level of commitment and responsibility that may translate to a reduction in violent behavior. The dynamic conceptualization provides a better theoretical and practical understanding of the relationship between the construct measured and violent recidivism than does the static approach.

Based on a dynamic conceptualization of risk of violence, the VRS (Part A) rating system redefines a number of accepted static measures as dynamic constructs. Perhaps most notable of these is "Criminal Personality" (item 2), which is similar to the construct of "psychopathic personality" as described by Factor 1 in the PCL-R literature. The "Criminal Personality" item on the VRS, although obviously based on Hare's (1991) notion of psychopathy, conceptualizes many psychopathic behaviors as amenable to change. For example, feelings of grandiosity, a tendency to manipulate others, and superficial presentation are a few traits/behaviors that could be susceptible to change through treatment or other life experiences. The assumption that psychopathic

behaviour is persistent across the life span, especially in terms of its relationship to violence, has not been supported empirically (Harpur & Hare, 1994).

All 26 items of the VRS (Part A) are each rated as 0, 1, 2, or 3; representing the degree to which that item relates to violent behavior for the offender in question. Therefore, two offenders could hypothetically be at very similar overall risk levels, as measured by the VRS total score, but have very different risk areas, as measured by individual dynamic risk items.

2.2.2.2 VRS Ratings of Change

The VRS is designed to evaluate changes on the set of 20 dynamic risk items. This design is ideal for a correctional treatment setting, in that many offenders may be able to reduce their risk through gains made in correctional treatment.

Ratings of change in risk are assessed by the VRS (Part B) through a 4 point rating reduction scale consisting of 0, -0.5, -1, -1.5. A value of "0" for a VRS risk item would indicate that no gains have been made to reduce risk for that risk item. A score of -1.5 would indicate that a substantial change has occurred during the treatment period and should result in a significant reduction in risk on that item. Ratings of -0.5 and -1.0 reflect intermediate levels of change equally spaced between 0 and -1.5. All VRS (Part B) item ratings use VRS pre treatment item ratings (Part A) as a baseline. The sum of the VRS (Part A) rating and the VRS (Part B) rating provides a derived post treatment risk score (VRS post treatment score). This process produces a derived post treatment rating which is not independent of pre treatment ratings or evaluations of change.

It is important to recognize that offenders must make therapeutic gains across several VRS risk items before overall risk levels would be expected to decrease. This is consistent with the notion that treatment gains must occur across different life domains

before therapeutic gains can have a positive impact on violent risk. In addition, risk reduction on individual items should provide a basis to better understand overall changes in risk. Individual VRS items provide qualitative and quantitative information as to where specific treatment changes have occurred (Toni et al., 1999). This gives the VRS an advantage over other risk measures in terms of understanding changes in offender risk and examining the impact of treatment. In short, the VRS not only has the capacity to provide a more sensitive evaluation of risk than more static based measures, but may be potentially employed to evaluate the effectiveness of correctional treatment.

2.3 Comparing the Predictive Efficacy of Risk Instruments

As stated earlier, as a result of the very low base rate of violence, even among criminally violent populations, the predictive accuracy of risk assessment instruments has been limited. In particular, incidents that occur very infrequently present a serious obstacle to predictive models. In particular, the future occurrence of rare events is generally over predicted (Meehl, 1954). In order to better understand errors in risk prediction it is helpful to conceptualize risk decisions in terms of positive decisions (i.e., decisions which predict a particular future event will occur) and negative decisions (i.e., decisions which predict a particular future event will not occur).

In essence, errors committed in predicting events with very low base rates most likely will be false positive, that is, the event predicted to occur in the future, does not occur. The other form of error, that is the commission of a false negative error, is the occurrence of an event when it was predicted that the event would not occur. False positive and false negative errors are central in determining the utility of a psychological judgements and assessment instruments. The concepts of test specificity and test

sensitivity allow for the examination of the false positive and false negative error rates when evaluating the utility of a risk prediction instrument.

A test is considered to have good specificity if the instrument can correctly identify a large percentage of the cases in which the specified outcome does not occur. In risk predictions, test specificity is an instrument's ability to correctly identify all low risk offenders (i.e., those who do not recidivate) by avoiding the error of categorizing a "real" low risk offender as a high risk offender. In essence, test specificity evaluates the ability of an instrument to avoid false positive errors (i.e., specificity = 1- false positive rate).

A test is considered to have good sensitivity if the instrument can correctly identify a large percentage of the cases in which the specified outcome does occur (i.e., sensitivity = the true positive rate). In terms of risk prediction, test sensitivity is an instrument's ability to correctly identify high risk offenders, and to avoid erroneously categorizing a "real" high risk offender as a low risk offender. In short, test sensitivity evaluates the ability of an instrument to avoid false negative errors. False negative errors are very unlikely when predicting events that have a low base rate because the actual occurrence of the predicted event is very infrequent to begin with.

Typically, professionals who predict events with low base rates, such as violent recidivism, are highly susceptible to false positive errors. However, it is important to recognize that the violent recidivism rate associated with samples of high risk offenders may be significantly higher than the rates associated with more typical forensic samples. For example, in risk predictions with offenders who have entered into high intensity treatment for violent behavior, there may be a relatively higher violent recidivism rate, and clinicians may need to be wary of making both "false negative" errors and "false

positive" errors. For this reason, some researchers have suggested that the focus of assessing the predictive efficacy of risk measures should be on "true" versus "false" positive error rate comparisons (Mossman, 1994). This suggests that an evaluation of test specificity and sensitivity should be the central focus in judging the utility of risk prediction instruments.

The Receiver Operating Characteristic (ROC) method of analysis is a convenient way to depict the ratio of "true" versus "false" positive rates (Mossman, 1994). In addition, researchers can avoid the difficulties normally associated with low base rate events by using a comparison of "true" versus "false" positive rates. The impact of a low base rate is already inherent to the "true" versus "false" positive ratio and, therefore, the inclusion of the base rate into the ROC analysis is not necessary. ROC analysis has a substantial advantage over other predictive tests in that it utilizes both instrument specificity and sensitivity to evaluate the predictive utility of assessment measures.

The ROC method can provide a comparison of predictive accuracy in terms of how an instrument may perform in relation to "chance" judgements, and in comparison to other risk measures. It is important to note that predictive accuracy of a risk instrument in comparison to "chance" is evaluated in terms of test specificity and test sensitivity. "Chance" is defined specifically as the case in which there is an equal proportion of true and false positive decisions (Mossman, 1994). That is, given that it is predicted that an offender will recidivate violently, the accuracy of this prediction is incorrect in exactly half of the cases. This would indicate that the specificity (i.e., 1 - the false positive rate) of the instrument was 50%, that is, in half of the cases in which violent recidivism is predicted, there are no subsequent violent convictions. Conversely,

in exactly half of the predicted cases, subsequent violent re-convictions did occur. This would indicate that the sensitivity of the instrument, or the true positive rate, is also 50%. This hypothetical situation represents "chance", in that the predicted positive outcomes were correct in exactly half of the cases, and that the number of incorrect positive decisions misidentified exactly half of offenders who did not recidivate.

A ROC curve is plotted with the vertical axis representing the percentage of true positives (i.e., test sensitivity), and the horizontal axis is plotted representing the percentage of false positives (i.e., 1- test specificity). By plotting the true positive rate (TP) against the false positive rate (FP), ROC analysis provides a method to incorporate and simultaneously evaluate the specificity and sensitivity of risk instruments. A line running through the origin of the axis and the subsequent ratio point provides a method to determine the predictive strength of the risk instrument. A steep line would represent a relatively higher true positive rate with a relatively lower false positive rate. A point in which these rates were equal or at "chance" (i.e., TP=.5, FP=.5), would be at 45 degrees. Any line at an angle greater than 45 degrees would be an improvement over chance, and conversely, any angle less than 45 degrees would represent predictions that are at less than chance. These three types of outcomes are represented in Figure 2.1.

Using the ROC method, Mossman (1994) indicated that the area under the ROC curve / line (AUC) can be used to summarize the overall discriminating power of the instrument. In the extreme case, for a line which is very steep (i.e., the rate of true positives is much greater than the rate false positives) the AUC would be close to 1.00, and for a line which has a very shallow slope (i.e., the rate of false positives is much greater than the rate of true positives) the AUC would approach 0. Mossman demonstrated that the AUC can be used to provide a meaningful comparison of

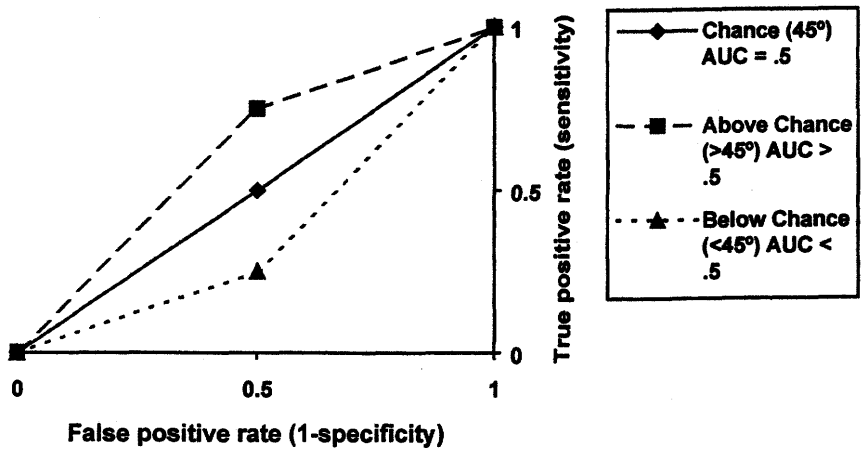

Figure 2.1 ROC Analysis of Predictive Performance at Chance, Above Chance and Below Chance Levels

predictive variables / instruments to chance levels (i.e., AUC= .5) or to provide a means to compare variables / instruments to one another. For example, the AUC values of first generation risk assessments can be compared to the AUC values of second generation risk instruments as measured by AUC values.

However, it is important to note that for ROC analyses, a scale must use a "cut off" or decision threshold score to predict an outcome. A specific rating value must be determined to indicate when the scale predicts the occurrence of the event in question in order that true and false positive rates can be calculated.

Rice and Harris (1995) have indicated that ROC analysis is of most practical value when several data points or decision thresholds are available, in order that the "curve" of the analysis is not represented by a straight line. Rice and Harris supported this assertion empirically by using analyses which set four and five decision thresholds or "cut off" points. It seems that ROC analysis based on multiple "cut off" points provides a better understanding of the predictive utility of an instrument, theoretically and empirically.

2.4 Objectives of the Present Study

The purpose of the present study was to investigate the predictive utility of the VRS with a sample of violent offenders who have completed at least four months of an intensive anger management treatment program. Specifically, the first objective of the study was to assess the ability of the VRS to measure changes in risk based on treatment information obtained from clinical case files. The second objective of the study was to compare the predictive efficacy of the VRS and the PCL-R.

To address the first objective of the study, the relationship between changes in risk as assessed by the VRS (Part B) and changes in the pattern of violent offending was examined. In particular, reductions in criminal offending patterns after treatment were expected to be correlated with a reduction in risk as assessed by the VRS (Part B). In addition, the assessment of the predictive utility of change (Part B) was also assessed by comparing VRS post treatment ratings and VRS pre treatment ratings. Presumably, if the assessment of change is important in predicting violent recidivism, then VRS post treatment ratings should demonstrate a stronger relationship with, and be more predictive of, violent recidivism than VRS pretreatment ratings.

The second objective of the study was addressed through three comparisons of the VRS with the PCL-R. First, a comparison of the strength of the relationship of the VRS pre and post treatment ratings and PCL-R ratings with violent recidivism was investigated. Second, VRS post treatment ratings and PCL-R ratings were compared in their ability to predict violent recidivism. In particular, VRS post treatment ratings were evaluated as to whether they provided a unique and significant contribution over PCL-R ratings in the prediction of violent recidivism. Finally, VRS pre and post treatment ratings and PCL-R ratings were compared in their relative efficacy to predict violence using ROC analysis.

Considering that treatment effects may diminish over time, and/or that VRS dynamic variables may be susceptible to further change after treatment, all of the analysis was conducted at a 1, 2, 3, 4, and five year follow up times, as well as a total follow up time. This process may assist in determining both the influence of treatment and the influence of time on the accuracy of risk predictions.

3. METHOD

3.1 Sample

3.1.1. Sample Selection

The present sample (n = 60) was selected on the basis of four criteria. In order to evaluate changes in risk levels, all offenders in the sample had participated, for a minimum of four months, in a treatment program which focussed on the reduction of violent behaviour. Second, all offenders had completed treatment between 1981 and 1992, thereby providing a sufficient follow up time (M = 90.03 months; SD = 32.04) to assess violent recidivism rates. Third, the sample was rated using the PCL / PCL-R during their treatment admission. Previous PCL / PCL-R ratings were used for the current study to provide a comparison measure for VRS ratings, and to provide a sample which reflected a broad range of offenders at high, medium and low levels of risk. To ensure that a broad range of offenders are represented in this study, the treatment consisted of an equal number (n=20) of high (PCL-R rating $\geq$ 25), moderate (17 $\leq$ PCL-R rating $\geq$ 24), and low (PCL-R rating $\leq$ 16) risk offenders.

There were a total of 155 federal offenders who had received PCL / PCL-R ratings at RPC, completed treatment during the specified time, and were involved in treatment for a minimum of four months. Using PCL-R ratings as a risk measure, there were only 20 offenders within the sample of 155 federal offenders who met the definition of a low risk offender (i.e., PCL-R rating $\leq$ 16). These low risk offenders comprised the low risk offender group in the present sample. In order to establish an equal number of offenders at each risk level, twenty medium and twenty high risk

offenders were selected to represent the medium and high risk groups for the present sample. For the moderate and high risk groups, as much as possible, the most recent offender case files were selected for rating purposes. Specifically, for the moderate and high risk offenders, treatment files were generally selected from the among the recent files available. This decision was based on observations that more recent case treatment files contained more comprehensive and relevant information for VRS coding.

3.1.2 Sample Characteristics

The present sample consists of 60 federal offenders treated in the violent offender program at the Regional Psychiatric Center (RPC). As indicated previously, all offenders within the sample had participated in a correctional treatment program for a minimum of four months. In addition, the sample was comprised of an equal number of high (n=20; PCL-R rating $\geq$ 25), moderate (n=20; 17 $\leq$ PCL-R rating $\geq$ 24), and low (n=20; PCL-R rating $\leq$ 16) risk offenders as identified by the PCL-R..

The three risk groups did not differ significantly in terms of total follow up time since release, age at time of release, length of time between date of first conviction and date of admission into treatment, but did differ significantly in terms of time spent in treatment (F (2, 57) = 6.78; p = .002), with the low risk group (M = 11.13 months; SD = 4.6) spending more time in treatment than either the high (M = 8.3 months; SD = 3.5) or medium (M = 7.15 months; SD = 2.22) risk groups.

3.1.3 Treatment Setting

The Regional Psychiatric Center (RPC) in Saskatoon, Canada provides progressive correctional treatment specifically targeting violent behavior. Although treatment at RPC has evolved in various respects over the past few decades, a strong

cognitive-behavioural approach utilizing anger management skills training techniques has been relatively consistent.

Treatment files for patients at RPC include admission and discharge reports, weekly progress evaluations, and nursing notes which are used to record significant daily events and the progress in treatment of each offender during the program. The RPC treatment files were the basis of VRS ratings for the present study.

3.2 Procedure

3.2.1 Psychopathy Ratings (PCL / PCL-R)

All offenders in the treatment sample were rated using the PCL or PCL-R by trained RPC research staff. Ratings were based on interviews and file information. Ratings were made at, or shortly after, an offender's admission into the treatment program. Ratings made by staff using the PCL were converted to PCL-R ratings. This conversion allowed for more consistent comparisons to be made across offender groups.

3.2.2 Recidivism Data

All offenders in the sample completed treatment at least seven years ago, and had been released back into the community for at least six months (only one case had less than 18 months follow up) with an average follow up time of 7.5 years. Follow up data (i.e., recidivism data) were obtained through the Offender Management System (OMS). OMS is a database used by Corrections Canada to collect information and monitor the progress of all federal offenders. One source of information utilized by the OMS database is the Canadian Police Information Service (CPIC), which is a list of all charges and convictions of Canadian federal offenders, the dates of any convictions, the charges associated with each conviction, and the disposition imposed by the court for

these convictions. The CPIC pertaining to each offender in the sample was utilized in this study to obtain recidivism data.

Although official recidivism is an underestimation of violent offending, in that many violent offenses may not be reported or successfully prosecuted, recidivism data does offer several advantages over other types of outcome measures. Official violent conviction data presents a relatively objective, clearly defined, and reliable method of assessing violent offending. In addition, one might expect that violent recidivism may be less of an underestimate of more serious violent offending, in that a higher intensity of violence may be associated with higher rates of reporting and detection, and with increased pressure for more successful prosecution.

3.2.3 Measures of Recidivism

Violent recidivism after treatment was measured by three methods. Cumulative number of post treatment violent convictions was the first measure of violent recidivism. This measure assessed the raw total number of violent convictions after treatment. Cumulative number of post treatment violent convictions was examined at 1, 2, 3, 4 and 5 years after treatment. In addition, number of violent convictions at the total follow up time for each offender also was examined. Cumulative rate of violent convictions was the second measure of violent recidivism. Rate of violent convictions was calculated by dividing the number of cumulative post treatment violent convictions by the number of months within the respective follow up time period. For example, the number of violent convictions in the first year after release was divided by 12, in order to provide a monthly rate of violent convictions for the first year after release. The cumulative number of violent convictions within the first two years was divided by 24, in order to provide a monthly rate of violent convictions within the first two years. The current

study calculated the rate of violent offending within 1, 2, 3, 4, and 5 years after treatment, and at total follow up time. The third measure of violent recidivism was focussed on the presence or absence of violent re-convictions in the follow up time period. An offender who had not recidivated with a violent crime would be coded "0" (i.e., absence), and an offender who has recidivated with one or more violent convictions would be coded "1" (presence). Presence or absence was examined at 1, 2, 3, 4, and 5 years after treatment, and at total follow up time.

3.2.4 Measures of Change

Changes in violent offending patterns were assessed through comparing the pre treatment rate of violent convictions with the post treatment rate of violent convictions. In order to calculate the rate of violent offending before treatment, the total number of violent convictions an offender had received before treatment was divided by the total number of months between the date of the offender's first violent conviction and his date of admission into the treatment program. Therefore, this rate was considered to be the average monthly violent conviction rate before treatment (VCR pretreatment). In order to calculate the rate of violent offending after treatment, the total number of violent convictions an offender had received after his discharge from treatment was divided by the total number of months between the date of treatment discharge and the date of data collection (i.e., 01/ 01 /99). This rate was considered to be the average monthly violent conviction rate after treatment (VCR post treatment).

Changes in risk levels were rated through the VRS (Part B) using the VRS coding manual. Although this section was developed primarily as a method to calculate post treatment risk levels, it also can be utilized as an independent evaluation of change (Wong & Gordon, 1996).

3.2.5 Information and Guidelines used for VRS Coding

All the data required for VRS ratings was obtained from the information contained in the treatment files at RPC. All VRS ratings were completed with raters blind to PCL-R ratings.

Although the same rater completed both the VRS pre treatment rating and the VRS post treatment rating for each offender file, VRS pre treatment ratings were completed before information concerning the treatment period was examined to ensure that knowledge of treatment progress did not bias pretreatment coding. This method follows the suggested guidelines of the VRS manual (Wong & Gordon, 1996) and was considered to be the method which most closely resembled the real life application of the VRS. Information examined during the treatment period was utilized only for post treatment VRS ratings and, therefore, was not used to adjust pre treatment ratings. This decision was based on attempts to maintain a reliable coding method, and in order to be consistent with the real life application of the VRS.

Ratings were completed in accordance with the third draft edition of the VRS manual which was completed in October of 1998. Although subsequent revisions and alterations have been made to the treatment change portion of the manual (Part B) since this date, the aforementioned version was determined to be adequate for the coding purposes of this study. The pretreatment rating system (Part A) has remained relatively unchanged to this point in time. The next completed version of the VRS coding manual is expected to be finished by the fall of 1999.

The current study relied on file information to determine ratings. The semi-structured interview suggested in the draft VRS manual was not incorporated into the proposed study because of the study's retrospective nature. At this time, research into

the comparison of VRS ratings derived from the interview and file information have not been compared to VRS ratings derived from only file information. In particular, this study may help to evaluate the utility of using only file information for future VRS research.

3.2.6 Inter-rater Reliability of VRS Ratings

A total of four raters, including the primary investigator, completed VRS pre treatment and VRS post treatment ratings for the 60 offenders in the treatment sample. Inter-rater reliability was established through a comparison of each coders' ratings with the ratings of the primary investigator. Initially, five files coded at the beginning of the coding process were used for calculating inter-rater reliability. Two of these initial files were "practice" files, and were not part of the current sample. The remaining three files were the first "real" files which were used in the present study. In addition, at the completion of the entire coding process, another three files from each rater were selected randomly by the primary investigator. That is, of the unique set of files each coder had rated at the conclusion of data collection, the primary investigator selected three files, which he subsequently rated. These additional three files were incorporated in determining the inter-rater reliability for each rater. Therefore inter-rater reliability was calculated based on first five files each coder rated, and on a random selection of three files that the coder had subsequently rated.

Inter-rater item reliability was assessed based on the total number of items which were coded. Each dynamic pre treatment VRS item is coded on a 4 point scale, and each dynamic post treatment adjustment is coded on a 3 point scale. Consistent with past VRS inter-rater reliability research (Wong & Gordon, 1996), the current study used a criteria of at least 60% of item ratings to be an exact match with the ratings of the

primary investigator, and at least 95 % of items to be an exact match or be only 1 point discrepant. This criteria was met in the current study. Inter-rater reliability of the three coders, based on the comparison using eight files for exact matches was 63%, 65%, and 76% for Part A and 64%, 78%, and 86% for Part B. Inter-rater reliability for an exact match or only 1 point discrepant was 99%, 97%, and 98% for Part A, and 97%, 97%, and 99% for Part B.

The inter-rater reliability of the six static items of the VRS was not assessed as they either are completely objective (e.g., age of first violent conviction), or they have very limited subjectivity due to the guidelines provided in the VRS manual (e.g., violence throughout the life span).

In terms of total score values for the VRS, the mean difference between each rater and the primary investigator was calculated for both VRS Part A total scores and for VRS Part B total scores based on the eight reliability files. Mean differences and standard deviations for total scores for VRS Part A (i.e., $\underline{M} = 3.75$, $\underline{SD} = 3.01$; $\underline{M} = 4.25$, $\underline{SD} = 3.69$; $\underline{M} = 4.75$, $\underline{SD} = 2.19$), were relatively small. Mean differences and standard deviations for the total scores for VRS Part B (i.e., $\underline{M} = 1.44$, $\underline{SD} = 0.90$; $\underline{M} = 1.88$, $\underline{SD} = 2.82$; $\underline{M} = 0.63$; $\underline{SD} = .64$, respectively) were also considered relatively small.

The internal reliability of the VRS (based on 26 items), as measured by Cronbach Alpha, was considered strong (i.e., Alpha = .88). The current sample had an approximately normal distribution of VRS ratings ($\underline{M} = 49.1$; $\underline{SD} = 11.5$).

3.2.7 Determination of Decision Points for ROC Analysis

ROC analysis requires that a cut off score must be established in order to calculate the false positive rate (i.e., 1 - test specificity) and the true positive rate (test

sensitivity) of the test instrument (Mossman, 1994). That is, there must be a rating value which is used to predict that an offender is likely to recidivate violently. A minimum of three values are necessary to develop an ROC curve, and to subsequently calculate an AUC value for the test instrument. For the present study, five decision cut off points were used to determine the ROC curve for each test instrument. In addition, the five decision points were distributed approximately evenly across the range of rating values so as to provide the best ROC curve estimate (i.e., low, medium, and high cut off rating values were selected from the distribution of ratings) for each test instrument. The same cut off values were used for the VRS pre treatment and post treatment rating scales for comparison purposes.

The first two cut off values for the VRS (i.e., 46 and 55) were selected by dividing the sample into three risk groups based on VRS ratings. As no criteria points have been established for determining risk groups based on VRS ratings, the sample was divided into three equal sized groups (i.e., n=20 for each group) based on the magnitude of VRS pre treatment ratings. Therefore, the lower "cut off" point for the middle group (i.e., VRS value of 46) and the lower "cut off" point for the high group (i.e., VRS value of 55) were used as decision "cut off" points for the VRS ROC analysis. In addition, a value between these two "cut off" points (i.e., 50) was used in the analysis as a the third decision point. Finally, "cut off" points based on identifying the highest 10 VRS ratings (i.e., 60) and the 10 lowest VRS (i.e., 40) were used as the fourth and fifth cut off points. It is expected that these latter two extreme values may provide an estimate of the floor and ceiling effect of the VRS in terms of its ability to predict violence. Therefore, ROC analysis of VRS pre and post treatment ratings are based on five threshold decision points (i.e., 40, 46, 50, 55 and 60). Admittedly, the determination process for these

decision points was somewhat arbitrary, however, it is hoped these decision points may provide some insight for future research in terms of determining offender risk groups based on VRS ratings.

To establish consistency, the ROC analysis using the PCL-R also utilized five "cut off points" using the same selection process as utilized for the determination of VRS "cut off" points. The PCL-R "cut off" values for ROC analysis were 13, 17, 24, 27 and 30.

4. RESULTS

4.1 Analysis of Change

4.1.1 Comparison of Pre and Post Treatment Number and Rate of Violent Convictions

There were significantly fewer violent convictions (t (59) = -6.76, $p < .001$) in the post treatment period (M = 1.47; SD = 1.85) than in the pre treatment period (M = 5.17; SD = 4.3). In addition, the mean monthly rate of violent convictions was also found to be significantly smaller (t (59) = -4.69; $p < .001$) for the post treatment period (M = .02; SD = .02) than for the pretreatment period (M = .08; SD = .11).

4.1.2 Comparison of Pre and Post Treatment VRS Ratings

Compared to VRS pre treatment ratings (M = 49.22; SD = 10.77), there was a significant reduction (t (59) = -9.13, $p < .001$) in VRS ratings after treatment (M = 46.83; SD = 10.57). Although this difference was statistically significant, it should be noted that mean change of VRS ratings (i.e., VRS Part B ratings) was relatively small (M = 2.4).

4.1.3 Correlation Between VRS Ratings of Change (Part B) and Changes in Post Treatment Number / Rate of Violent Convictions.

There was no significant Pearson correlation between VRS Ratings of Change (Part B) with the change in number of violent convictions between pre and post treatment. There was no significant Pearson correlation between VRS Ratings of Change (Part B) with the change in rate of violent convictions between pre and post treatment.

4.2 Comparison of PCL-R and VRS with Three Measures of Violent Recidivism

The correlations between PCL-R, VRS pretreatment and VRS post treatment ratings with cumulative number of post treatment violent offenses, cumulative rate of post treatment violent offending (Pearson r), and presence / absence (point biserial) of violent recidivism are presented in Table 4.1.

Considering the small sample size in the present study, and the small magnitude of correlations generally found in research predicting violent recidivism (e.g., .25-.35; Harris et al., 1993), the present analysis considered differences in correlation coefficients of 0.1 to reflect a meaningful difference. For example, if risk instrument "A" correlated with the number of post treatment violent offenses at .28, then the correlation of risk instrument "B" must equal or exceed .38 (i.e., .28 + .1 = .38) in order to provide a meaningful advantage.

As illustrated in Table 4.1, VRS pre and post treatment ratings did not differ in terms of their relationship with violent recidivism at any of the time intervals. In particular, VRS post treatment ratings did not provide a correlational advantage over VRS pre treatment ratings. Due to this similarity, further discussion of correlations with recidivism will use the term "VRS ratings" to refer to both VRS pre and post treatment ratings.

As also illustrated in Table 4.1, PCL-R ratings provide a correlational advantage over VRS ratings in the first year after release in terms of both cumulative number of violent convictions and in terms of cumulative rate of violent convictions. Conversely, VRS ratings appear to provide a correlational advantage over PCL-R ratings at the total follow up time for cumulative number and rate of post treatment violent convictions. At all other follow up times, all three measures were generally equivalent in terms of their

Table 4.1

Pearson And Point Biserial Correlation Matrix of PCL-R Ratings, VRS pre treatment ratings and VRS post treatment ratings with Number of Violent Convictions, Rate of Violent Convictions, and Presence or Absence of Violent Convictions.

	Number of Years After Release																	
Follow up time	1 Year post release (n = 60)			2 Year post release (n = 58)			3 Year post release (n = 58)			4 Year post release (n = 57)			5 Year post release (n = 54)			Total follow up time (n = 60)		
Recidivism Measures / Risk Instrument	NUMBER	RATE	YES/NO	NUMBER	RATE	YES/NO	NUMBER	RATE	YES/NO	NUMBER	RATE	YES/NO	NUMBER	RATE	YES/NO	NUMBER	RATE	YES/NO
PCL-R	.33 b	.31 a	.27 a	.38 b	.41 c	.35 b	.39 b	.39 b	.34 b	.37 b	.37 b	.27 a	.35 b	.35 b	.27 a	.29 a	.30 a	.14 ns
Pre treatment VRS	.22 ns	.22 ns	.27 a	.40 b	.34 b	.47 c	.36 b	.36 b	.34 b	.32 a	.32 a	.20 ns	.31 a	.31 a	.18 ns	.41 c	.40 c	.22 ns
Post treatment VRS	.24 ns	.24 ns	.28 a	.38 b	.34 b	.45 c	.34 b	.34 b	.34 b	.30 a	.30 a	.21 ns	.30 a	.30 a	.18 ns	.41 c	.39 b	.23 ns

Note. a = $p < .05$; b = $p < .01$; c = $p < .001$ (two tailed test). ns = not statistically significant ($p > .05$). NUMBER = the number of post treatment violent convictions. RATE = the rate of post treatment violent convictions. YES / NO = presence or absence of violent recidivism.

relationships with cumulative number and rate of post violent convictions.

VRS ratings provided a correlational advantage over PCL-R ratings at 2 years after release in terms of predicting the presence or absence of violent recidivism. However, overall, PCL-R ratings showed a more enduring relationship with presence/ absence of violent recidivism and provided a correlation advantage over VRS ratings at year 4 and year 5. Despite this advantage, PCL-R correlations at year 4 and year 5 were relatively weak, and the relationship between all three risk instruments and the presence/ absence of violent recidivism greatly diminished after 3 years post release.

In general, over 1-5 years after release, PCL-R ratings demonstrated a relatively stable relationship with cumulative number and rate of violent convictions and presence/ absence of violent recidivism, with a slight peak evident at year 2 and year 3. Similarly, VRS ratings showed a marked peak in their relationship with cumulative number and rate of violent convictions at year 2 and year 3, with a gradual decline after these points. Figures 4.1, 4.2, and 4.3 illustrate the pattern in the correlation coefficients between PCL-R and VRS ratings with cumulative number, rate, and presence/absence of violent convictions respectively.

4.2.1 Multiple Regression Using PCL-R and VRS Post Treatment Ratings in the Prediction of Violent Recidivism

The results of simple multiple regression using PCL-R rating and VRS post treatment rating to predict cumulative number of post violent convictions, cumulative rate of post violent convictions and presence/absence of post treatment violent convictions are presented in Table 4.2. In addition, Table 4.2 also includes the regression equations at 1, 2, 3, 4 and 5 year follow up time, and at total follow up time.

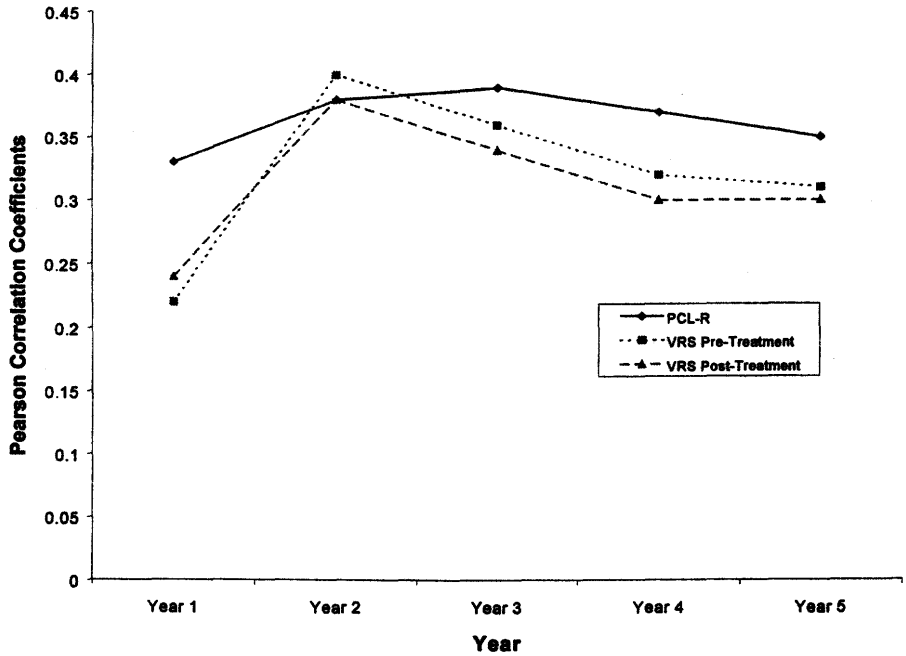

Figure 4.1 Pearson Correlations of PCL-R and VRS Ratings with Number of Convictions at 1-5 Years Post-Release

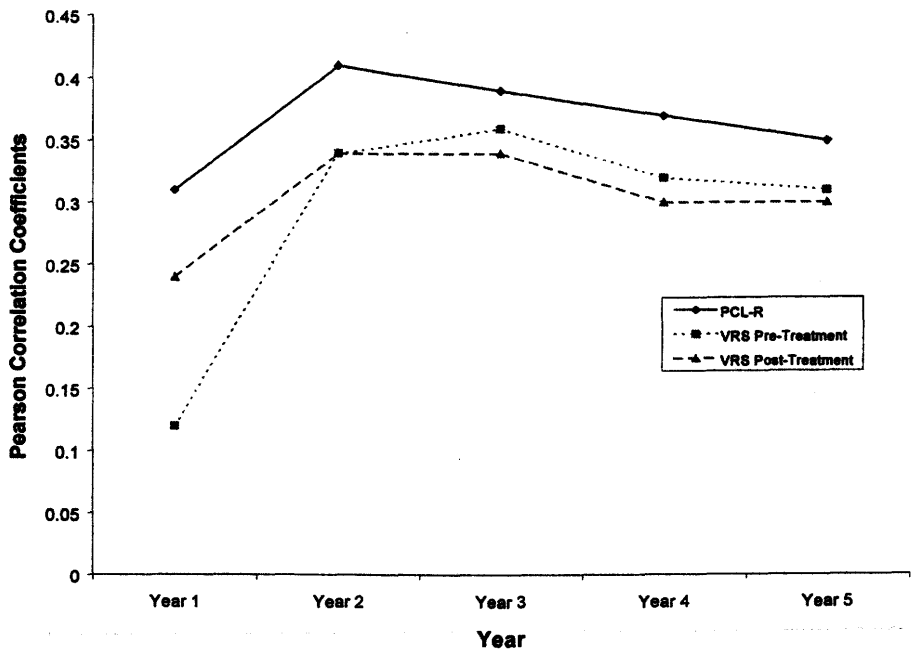

Figure 4.2 Pearson Correlations of PCL-R and VRS Ratings with Rate of Violent Convictions at 1-5 Years Post-Release

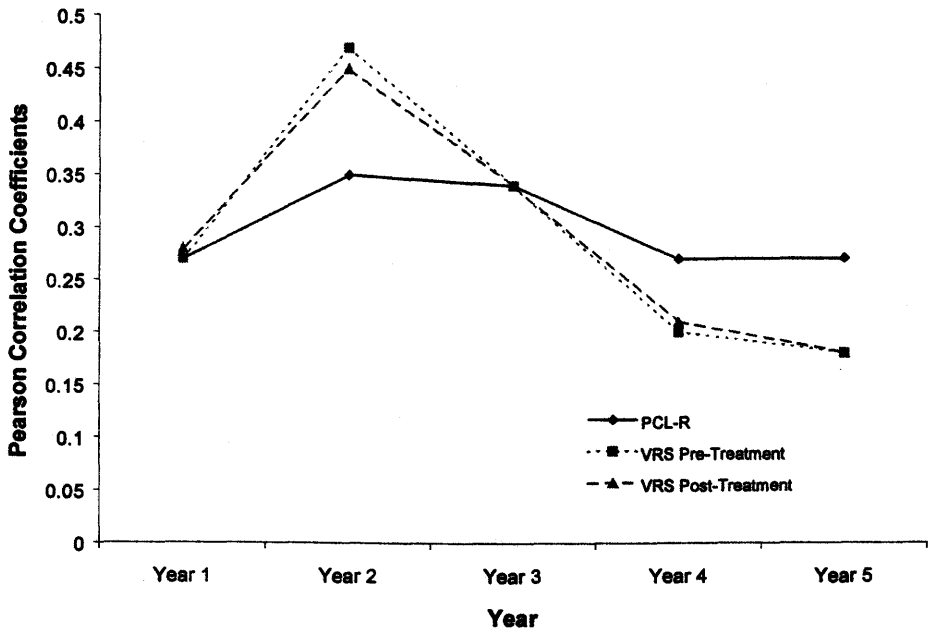

Figure 4.3 Point Biserial Correlations of PCL-R and VRS Ratings with Presence or Absence of Violent Convictions at 1-5 Years Post-Release

At one year after release, there were significant, but marginally predictive regression equations for all three measures of violent recidivism. PCL-R rating was the only independent significant predictor for cumulative number of violent convictions. PCL-R ratings and VRS post treatment ratings were not independently predictive in the regression equations for cumulative rate of violent convictions or for the presence/ absence of violent recidivism.

Within two years after release, regression equations were at their most predictive (i.e., as measured by Adjusted R square) for all three measures of violent recidivism. In terms of cumulative number of post treatment violent offenses, neither the PCL-R nor

Table 4.2

Multiple Regression Using PCL-R Ratings and VRS post treatment ratings to Predict Number of Violent Convictions, Rate of Violent Convictions, and Presence or Absence of Violent Convictions.

Follow up Time and Measures of Violent Recidivism	F Value of Multiple Regression Equation	Adjusted R square Value	t value of PCL-R in Multiple Regression Equation	t value of Post treatment VRS in Multiple Regression Equation
1 Year after release				
Number of VC	F = 3.84 (p = .027)	.09	t = 2.01 (p = .049)	t = .65; n.s.
Rate of VC	F = 3.17 (p = .050)	.07	t = 1.68; n.s.	t = .71; n.s.
Yes/ No	F = 3.18 (p = .049)	.07	t = 1.19; n.s.	t = 1.34; n.s.
2 Years after release				
Number of VC	F = 6.43 (p = .003)	.16	t = 1.80; n.s.	t = 1.76; n.s.
Rate of VC	F = 5.92 (p = .005)	.16	t = 2.19 (p = .033)	t = 1.08; n.s.
Yes/ No	F = 8.10 (p = .001)	.20	t = 1.28; n.s.	t = 2.66 (p = .01)
3 Years after release				
Number of VC	F = 5.96 (p = .004)	.15	t = 2.04 (p = .046)	t = 1.39; n.s.
Rate of VC	F = 5.96 (p = .004)	.15	t = 2.04 (p = .046)	t = 1.39; n.s.
Yes/ No	F = 4.98 (p = .010)	.12	t = 1.62; n.s.	t = 1.53; n.s.
4 Years after release				
Number of VC	F = 4.87 (p = .011)	.12	t = 2.01 (p = .049)	t = 1.08; n.s.
Rate of VC	F = 4.87 (p = .011)	.12	t = 2.01 (p = .049)	t = 1.08; n.s.
Yes/ No	F = 2.42; n.s.	.05		
5 Years after release				
Number of VC	F = 4.17 (p = .021)	.11	t = 1.76; n.s.	t = 1.06; n.s.
Rate of VC	F = 4.17 (p = .021)	.11	t = 1.76; n.s.	t = 1.06; n.s.
Yes/ No	F = 2.05; n.s.	.04		
Total Follow up time				
Number of VC	F = 6.09 (p = .004)	.14	t = 0.88; n.s.	t = 2.49 (p = .01)
Rate of VC	F = 5.74 (p = .005)	.15	t = 1.00; n.s.	t = 2.31 (p = .02)
Yes/ No	F = 1.65; n.s.	.02		

Note. VC = number of violent convictions. Yes/ No = presence or absence of violent recidivism. n.s. = not significant (p > .05)

the VRS post treatment ratings were significant independent predictors. For cumulative rate of violent convictions, the PCL-R was the only significant predictor. VRS post treatment rating was the only significant predictor of presence/ absence of violent recidivism.

Within three years after release, all three regression equations were significant. For both cumulative number of post treatment violent convictions and cumulative rate of post treatment violent convictions, PCL-R rating was the only significant independent predictor. Neither PCL-R rating nor VRS post treatment rating was a significant independent predictor of the presence/absence of violent recidivism.

Within years 4 and 5, and at total follow up time, there were only two significant regression equations for the prediction of cumulative number and rate of post treatment violent convictions. Within 4 years, only PCL-R rating was a significant predictor for cumulative number and rate of post treatment violent convictions. Within 5 years, neither PCL-R nor VRS post treatment rating was an independent predictor of cumulative number or rate of post treatment violent convictions. At total follow up time, only VRS post treatment rating was a significant predictor of number and rate of post treatment violent convictions.

In terms of the predictive strength of the regression equations over 1-5 years after release (i.e., as measured by the Adjusted R square value), the predictive ability of the multiple regression equations appeared to peak within year 2 and year 3 follow up, and gradually declined thereafter. Figure 4.4 illustrates the pattern of the Adjusted R square values of the multiple regression equation predicting cumulative number and rate of violent convictions within 1-5 years after release.

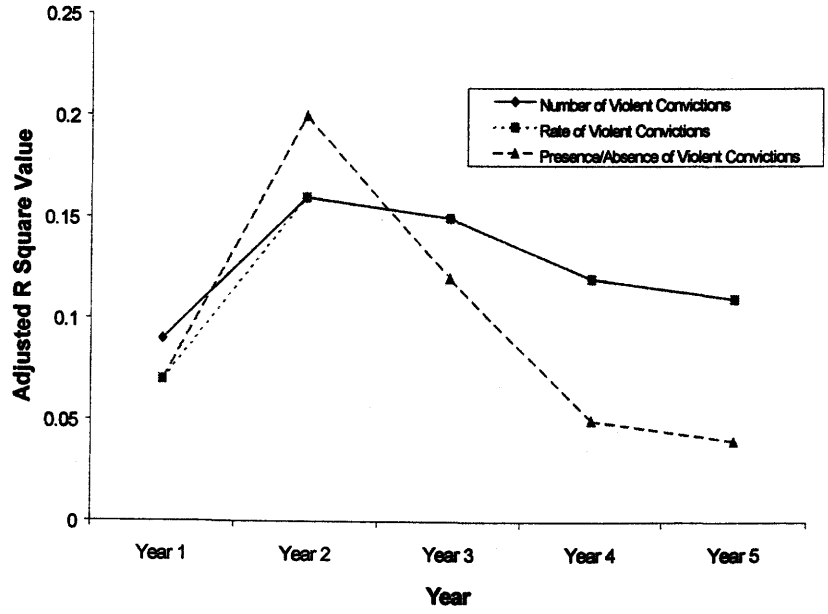

Figure 4.4 Adjusted R Square Values of Multiple Regression Using the PCL-R and VRS post treatment Ratings to Predict Number, Rate, and Presence/Absence of Violent Convictions

4.2.2 ROC Analysis

ROC curve analysis over 1-5 years and total follow up time for PCL-R, VRS pre treatment, and VRS post treatment are shown in Figures 1 - 18 in Appendix C. For comparison purposes, AUC differences between risk instruments of .05 or greater were used to determine an improvement of one measure in comparison to another. Although admittedly this criterion is arbitrary, for measures which perform above chance, a

Table 4.3

AUC Value Comparison Between the PCL-R, VRS pre treatment assessment and VRS post treatment assessment.

Risk Instrument	Number of Years After Release					
	1 Year	2 Year	3 Year	4 Year	5 Year	Total follow up time for all offenders
PCL-R	.64	.72	.67	.64	.63	.55
Pre treatment VRS	.66	.78	.70	.64	.61	.64
Post treatment VRS	.71	.80	.70	.62	.61	.61

difference of .05 represents approximately 10% of the available variance between tests, and, therefore, may be considered a substantial improvement.

At the second year following treatment, all three measures improved considerably, and were at their highest AUC value. In particular, the VRS pretreatment AUC value increased by .12. At the second year, both VRS pre treatment and VRS post treatment scores provided an advantage over the PCL-R in terms of the ROC analysis.

For the third, fourth, and fifth years, all three measures performed at comparable levels in terms of the AUC values. In particular, the AUC values for all three measures showed a decline after the second year. At total follow up time, the AUC of the PCL-R

dropped considerably below the AUC values of the VRS pre and post treatment AUC values and approached chance level (i.e., PCL-R AUC = .55).

In general, all three measures demonstrated a peak in their AUC values at year 2, with relatively more stable AUC values by years 4 and 5. Figure 4.5 illustrates the strength of AUC values of the PCL-R and VRS ratings at 1- 5 years post release.

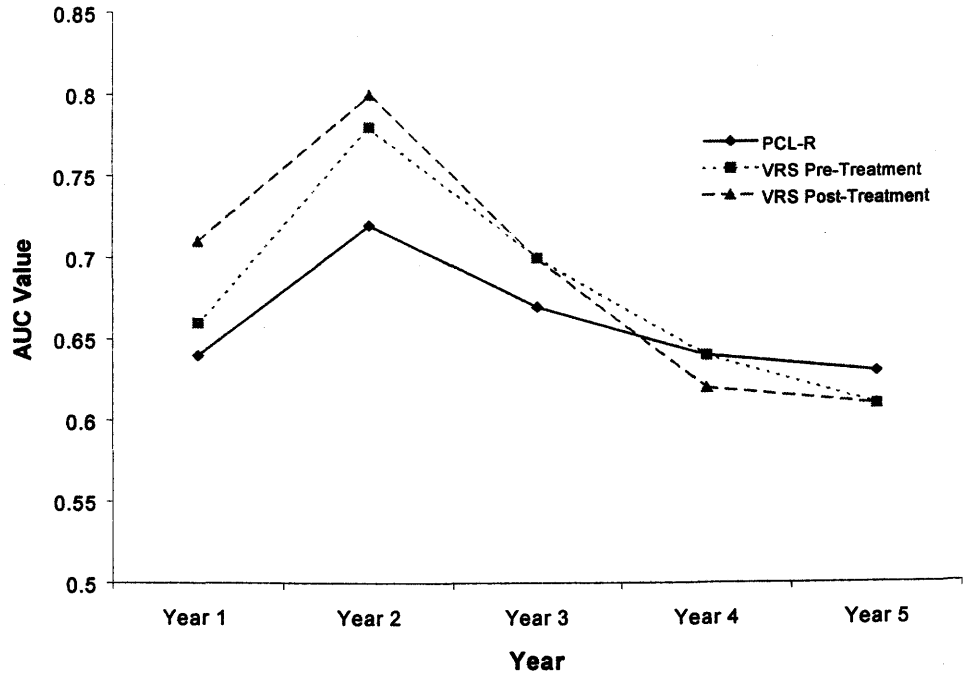

Figure 4.5 AUC Values of PCL-R and VRS Ratings at 1-5 Years Post Release

5. Discussion

5.1 Changes in Violent Offending and VRS Post Treatment Ratings

As illustrated by the significant reduction in the rate of violent offending after treatment, the present study demonstrates the importance of recognizing that violent offending patterns can change. As such, the ability of risk instruments to detect changes in risk levels is essential for risk prediction. Although the current study does illustrate the importance of changes in risk levels, it appears that VRS change scores did not significantly improve risk predictions and were not related to the reduction in violent offending. It may be that VRS ratings of Change (Part B) were too small in magnitude to improve the predictive ability of the VRS. Although the VRS post treatment scores were significantly lower than VRS pre treatment scores, the mean difference was relatively small, only about 2.5 points.

However, it is important to note that the AUC of the VRS post treatment ratings did offer an advantage over VRS pre treatment ratings at 1 year post release. Despite the general finding that VRS Ratings of Change were relatively small, they still provide some improvement to VRS ratings in terms of test sensitivity and test specificity, albeit only at a very immediate follow up time (i.e., the advantage is lost after the first year post release). This improvement is very important for decision makers who may be required to make decisions about risk within this one year window.

5.2 Limitations of VRS Ratings of Change

A few aspects of the present study may have limited the predictive utility of VRS ratings of Change (Part B). The current study is retrospective in nature, and therefore,

the inclusion of the semi-structured interview portion for VRS rating was not possible. The interview allows for an in depth exploration of specific high risk areas which is not available via file information alone. In addition, file information may be inadequate to assess changes in risk accurately. It was noted by raters that often a number of files, in particular, earlier treatment files, were not very comprehensive in describing the important areas of risk, provided insufficient collateral information, and were not very focussed on information about changes in risk. More recent treatment reports appeared to be better informed by contemporary research and have a better focus on risk areas which are targeted by the treatment program. A study using more contemporary file information to make VRS ratings may be better able to assess changes in risk in a way which relates to future violent offending patterns (Wong & Gordon, 1999).

Finally, the VRS coding manual used in the present study only provides for the measurement of a reduction in risk. In essence, a rater only rates risk as either stable or reduced. The revised version of the VRS coding manual also provides for the measurement of increases in risk level. Although in the current study raters indicated that increases in risk during treatment were relatively uncommon, a few offenders appeared to become more "criminalized" or "decompensated" on some risk items during the treatment period. In order to maximize the predictive validity of assessment of change in risk, a risk instrument must be able to reflect increases in risk.

Future research which addresses these shortcomings may better assess the utility of the VRS to measure changes in risk. Changes in offending patterns have been demonstrated to be measured effectively using the Criminal Career Profile (Wong, Templeman, Gu, Andre & Leis, 1997). This measure, although still in development, may be useful in future research investigating the utility of VRS ratings of Change (Part

B). In addition, the inclusion of other dynamic risk measures (e.g., the LSI-R, the HCR-20) also may be useful comparison measures to evaluate VRS ratings of Change.

5.3 Overall Patterns of the Current Findings: A Peak at Year 2

Due to the similarity in their relationship to post treatment violent convictions, VRS pre treatment ratings and VRS post treatment ratings most often will be collectively referred to as VRS ratings. In the case of specific comparisons, VRS post treatment ratings will be used.

The present study indicates that over a 5 year follow up period, the PCL-R and VRS ratings perform best at 2 to 3 years follow up. The pattern is relatively stable and is evident in correlation comparisons, the comparison of the predictive ability of regression models, and the comparison of AUC values derived from ROC analyses. In addition, this pattern is consistent whether recidivism is measured by cumulative number/rate of post treatment violent convictions or through the measurement of presence or absence of post treatment violent convictions.

The PCL-R demonstrates a more invariant pattern of results across time than does the VRS. Although the PCL-R generally performs best at year 2 and/or year 3 follow up, this improvement is generally modest. Conversely, the performance of the VRS appears to be influenced substantially by time: Year 2 and/or year 3 follow up provides a large improvement for the VRS. In particular, the post treatment violent convictions often is no longer significantly related to VRS ratings at follow up times which are longer than 3 years (e.g., correlation with presence / absence of violent convictions). In general, it appears that VRS ratings are superior or equal to PCL-R ratings at the 2 and 3 year follow up times, but the VRS performs less well after 3 years post release.

Consistent with this pattern, the ROC analysis indicates that all three measures have improved AUC values at year 2, and the improvement is largest for the VRS. At year 2, the AUC values of the VRS are significantly better than that of the PCL-R, but in years 3, 4 and 5, the three measures have approximately equal AUC values.

A similar pattern of results is also reflected in the predictive abilities (i.e., Adjusted R Square) of the regression equations using the PCL-R and VRS to predict cumulative number/rate and presence or absence of post treatment violent convictions at 1-5 years post release.

Given that the PCL-R has been used as a static measure of risk, it is not surprising that it is influenced less by the length of follow up time than is the VRS. Given its dynamic nature, the efficacy of the VRS would be expected to be a function of the length of follow up time; the longer the follow up time, the more the intervening variables (e.g., age, life situations, etc.) would dilute the VRS's predictive ability. Consistent with a static model of risk, the predictive efficacy of the PCL-R does not fluctuate very much over follow up time. The present finding indicates that the conceptual differences of the PCL-R and VRS are reflected empirically by the change in predictive power over time.

5.3.1 Static and Dynamic Risk Assessment

The "static" approach to risk assessment, as evidenced by the PCL-R, provides an unchanging measure of risk. This approach not only defines risk in a stable way, but as demonstrated by the findings in this study, is empirically related to risk in a relatively stable way, although it still appears to be influenced somewhat by time at 2 and/or 3 years follow up.

Conversely, the dynamic approach to risk assessment, as evidenced by the VRS, provides a measure of risk that can change in due course. This approach assumes that risk is changeable, and may be influenced by a variety of situational factors in an offender's life. Therefore, predictions based on dynamic risk factors will be effective within a limited follow up time. Analogous to a motion picture, a single dynamic risk assessment represents a single slide, which gives the assessor a "snapshot" of the important risk areas at that particular time. It is expected that the actual level of risk will gradually change due to new situational influences and, therefore, the "snapshot" will become more different from the actual risk level as length of follow up time increases. The present study suggests that the "life expectancy" of a dynamic risk prediction is about 2- 3 years. After this point, it is expected that intervening factors could have impacted various risk areas, and that an updated prediction or "snapshot" is necessary.

The present findings also suggest that comparisons between static and dynamic measures may need to be sensitive to the unique characteristics of these two approaches. For example, a comparison study with a follow up period of 2-3 years is likely to find that dynamic measures are superior to static measures. On the other hand, a comparison study using a 5 year follow up time likely will conclude that static measures are superior. It would seem that conducting traditional comparisons (i.e., comparisons which use a single follow up time; Hemphill, Hare & Wong, 1998) of risk instruments actually incorporates a static view of risk. Risk is seen as a relative constant on which measures can be compared to determine which matches most closely the stable "true" level of risk. The current study suggests that the "true" level of risk as assessed by dynamic measures will vary, depending on which time interval is of interest. The static

model is based on the premise that the time interval is of little consequence (although in fact, the longer the time interval, the more a static approach to risk prediction is advantageous). Thus, a comparison of two static instruments may lend itself more readily to the traditional comparison approach.

5.3.2 Improved Prediction at 2 Years Post Release

The increased performance of both the PCL-R and the VRS at about 2 years follow up may indicate that this is a time period at which there is the greatest discrimination in offending patterns between high and low risk offenders. This pattern may indicate that most high risk offenders who will recidivate violently will do so by the second year after release. In contrast, at this relatively short follow up time, most individuals who are considered low risk have not yet received any violent convictions. The results of post hoc analyses of the cumulative mean number of post treatment violent convictions for the high, moderate and low risk groups as classified by the PCL-R over 1- 5 years are presented in Figure 5.1. In addition, post hoc analysis of number of post release convictions was also conducted by dividing the sample into thirds based on the magnitude of VRS post treatment ratings. For this investigation, the upper third of VRS post treatment ratings were considered a high risk group, the middle third was considered a moderate risk group, and the lower third a low risk group. Post hoc analyses of cumulative mean number of post release violent convictions for the high, moderate and low risk groups as classified by the VRS over 1-5 years are presented in Figures 5.2.

As illustrated by Figures 5.1 and 5.2, high risk offenders, as identified by the PCL-R and VRS, demonstrate a substantial increase in their mean number of cumulative

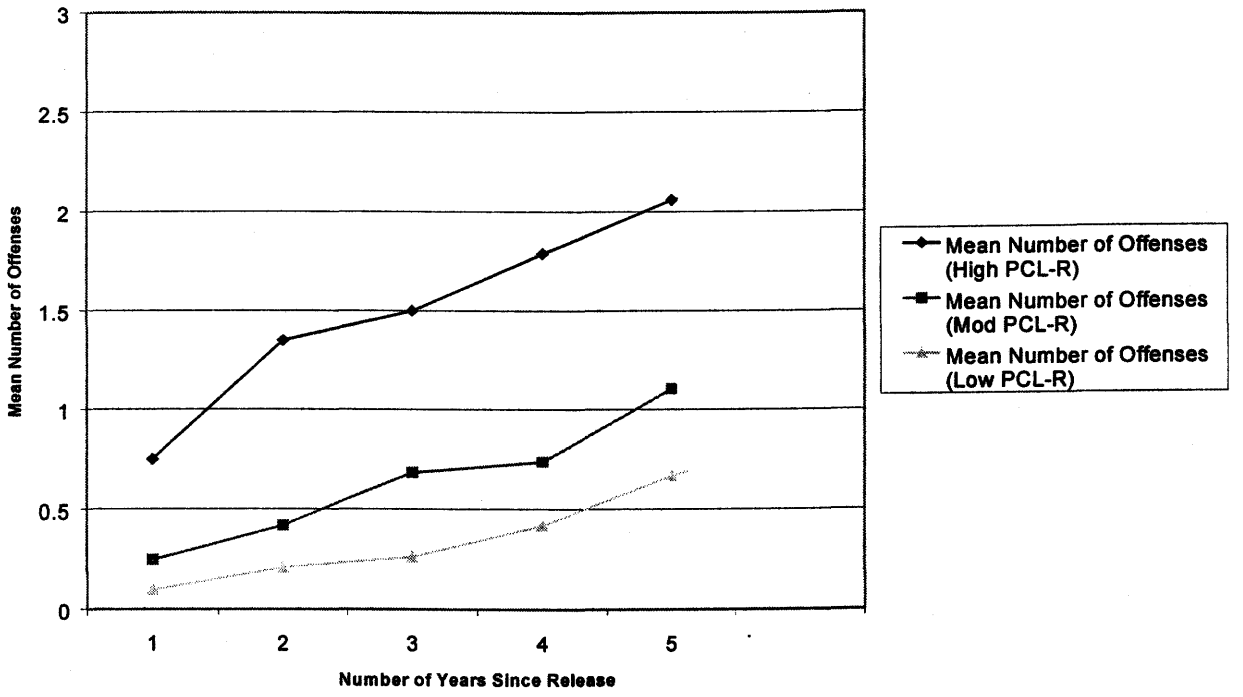

Figure 5.1 Cumulative Mean Number of Offenses by Risk Type (PCL-R)

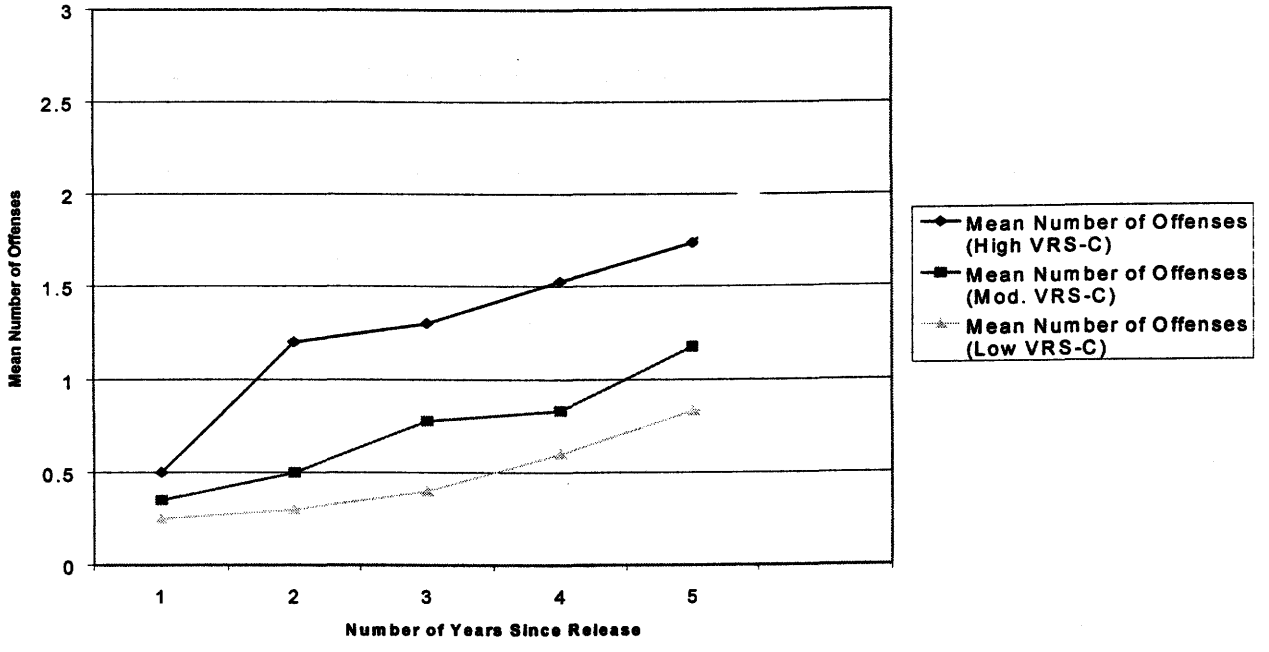

Figure 5.2 Cumulative Mean Number of Offenses by Risk Type (Post-Treatment VRS)

offenses at 2 years follow up. In addition, it appears that moderate and low risk offenders begin to increase their cumulative mean number of offenses after year 3. This finding is consistent with the notion that the second year post release marks a time period at which high and low risk offenders most differentiate themselves. Further support for this suggestion is provided in the examination of the respective percentages of high, moderate and low risk offenders who recidivate over the five year period. Percentage of offenders who have violently recidivated over the cumulative 1-5 year follow up times is presented in Appendix D. In essence, the majority of high risk offenders who will violently recidivate have done so by the 2 year follow up time, and a large portion of moderate and low risk offenders who will violently recidivate will do so after the two year follow up period. Therefore, the 2 year follow up time marks the period at which high and low risk offenders best differentiate themselves.

5.3.3 Pattern of Current Findings at Total Follow Up Time

Perhaps the most notable and consistent difference between the PCL-R and the VRS in their relationship with violent recidivism is found at the total follow up time period. The VRS provides a substantial advantage over the PCL-R at total follow up time in almost all of the present analyses.

Explanations for this finding are difficult due to the varied nature of the total follow up time variable. In general, follow up time varied greatly among the offenders in the current sample (M = 90.03 months, SD = 32.04 months); and ranges from approximately 6 months (only 1 offender had less than 17 months follow up) to approxiamately16 years (only 1 offender had more than 12 years follow up). Although 90% of the present sample had a follow up time which exceeded 5 years, the substantial variation in the sample provides an obstacle for interpretation of total follow up time.

However, despite this difficulty, total follow up time generally would seem to represent a varied extension of the 5 year follow up time.

A partial explanation of the improved relationship and predictive ability of VRS with violent recidivism at a longer follow up time may be the moderating effect of age. Clearly as follow up time increases, age becomes an increasingly important variable in assessing future risk. In particular, the empirical literature has demonstrated violent offending greatly decreases as offenders approach middle age (Harpur & Hare, 1994). "Current age" (i.e., age at assessment) is a static item measured by the VRS (but not by the PCL-R), and may provide some predictive improvement for the VRS as offenders in the sample become older. However, "current age" is only one of the VRS risk items, and this single item cannot account for the relatively high correlation of VRS ratings and number/ rate of violent convictions at total follow up time. A more comprehensive explanation is necessary to better understand the improved relationship to number/rate of violent convictions at total follow up time.

Unfortunately, because of the differences in follow up times of offenders in the present sample, the importance of total follow up time and its relationship to VRS ratings is unclear. Future research which systematically follows offenders beyond the five year period may address this finding. In particular, a year by year examination of the relationship between VRS ratings and number/rate of violent convictions over follow up times which exceed five years may provide a better indication as to when this relationship improves, and may better determine the source of this improvement. Unfortunately, due to the diversity of follow up time in the current sample, a systematic approach beyond the five year follow up was not practical (i.e., less than 80% of the

sample had data exceeding 6 years and under 60% of the sample had follow up data exceeding 7 years).

5.4 Comparison of the PCL-R and VRS

The Pearson correlation between the PCL-R and VRS was moderate ($r = .50$), suggesting that the two risk instruments were related, but not identical. This relationship is comparable with past comparisons of the PCL-R and VRS (Wong & Gordon, 1996).

As the post treatment rate of violent recidivism is derived from number of post treatment violent convictions, the following comparisons will include only the cumulative rate of violent convictions as this measure is less susceptible to differences in follow up time.

5.4.1 Comparison of the Relationship and Predictive Ability of the PCL-R and VRS with Rate of Violent Convictions

An important task for practitioners and researchers is to identify offenders who will habitually commit violent crimes at a high rate. Therefore, one important comparison for risk instruments is their ability to discriminate between high and low rate violent offenders. The present study indicates that although the correlations of the PCL-R ratings and VRS ratings with rate of violent convictions are approximately equal after the first year post release (i.e., the PCL-R has a correlational advantage at year 1), only the PCL-R provides a significant independent contribution in the prediction of rate of violent convictions in a simple regression model. Due to the finding that there were no predictive equations in which PCL-R ratings and VRS post treatment ratings were both significant, it is considered that these measures overlap in the risk information they utilize. Therefore, the PCL-R may be a preferred measure in this type of risk prediction.

However, it should be noted that the predictive abilities of the PCL-R for rate of violent convictions are present only within 2- 4 years follow up.

5.4.2 Comparison of the Relationship and Predictive Ability of the PCL-R and VRS in Determining the Presence or Absence of Violent Convictions

Another important task many correctional professionals are faced with is the prediction of whether an offender will commit any future violent offenses. In terms of answering this question, PCL-R ratings demonstrated a modest, but significant relationship to the presence or absence of violent convictions across 1 - 5 years post release. However, the relationship of the VRS with the presence or absence of violent convictions, although significant only at 1- 3 years after release, had a correlational advantage over PCL-R ratings at year 2. In terms of prediction, VRS ratings demonstrated the only significant contribution to the prediction of presence or absence of future violent convictions, which was at year 2. The current study also indicates a marked improvement in the predictive ability of the simple multiple regression equations (i.e., as measured by the Adjusted R square) at the second year after release, which diminishes in the third year, and is not significant by the fourth and fifth years.

More importantly, VRS ratings also demonstrated an AUC advantage over PCL-R ratings within the first two years after release. Because ROC analysis provides a means to assess both test sensitivity and test specificity, it provides a more thorough assessment of risk instruments than correlational analysis. In particular, the ROC analysis provides a basis to determine if the level of accuracy (i.e., true positives) of risk predictions are superior to levels of inaccuracy (i.e., false positives). The current findings indicate that VRS ratings provide a relatively high ratio of true positive decisions in comparison to false positive decisions within a two year follow up. At two

years after release, if decisions are set to a criterion of 95% true positive rate using VRS post treatment ratings, there is a corresponding 50% false positive rate. In comparison, at two years after release, if decisions are set to a criterion 95% true positive rate using PCL-R ratings, there is a corresponding 85% false positive rate. For the third, fourth and fifth years after release, the AUC values of the VRS and PCL-R were approximately equal.

The present study suggests that the VRS is preferable to the PCL-R in the prediction of presence or absence of post release violent convictions. As previously discussed, this advantage is most prevalent within the first 2 years of follow up and, after this time period, VRS ratings and PCL-R ratings are approximately equal.

5.4.3 Implications for Efficacy of the PCL-R and VRS

In general, the predictive efficacy of the PCL-R is more invariant over time than the VRS, which is consistent with the theoretical approaches of both measures. The PCL-R performs as well, or is more predictive than, the VRS in measuring cumulative number/rate of post treatment violent convictions across time, although the differences are generally small. This finding is consistent with the theoretical approach of the PCL-R, in that "psychopaths" are conceptualized to possess personality and behavioral features which make them more prone to committing violent offenses in a much more chronic and frequent manner than other offenders.

In terms of determining the presence or absence of violent recidivism, the VRS appears to be a superior measure, but only within the first 2 or 3 years after release. After this point the VRS is no longer as effective at predicting violent convictions. The finding that the VRS is related to future presence or absence of violent convictions is consistent with the design of the VRS. The VRS assesses offenders in terms of whether

each risk item is related to violent offending. A high risk offender is identified by having several "risk" items. Therefore, predictions are not necessarily based on expectations involving the frequency of future violence, but rather, focus on the likelihood of the offender to engage in any future violence. In essence, the VRS assesses the risk of an offender to behave violently based on their current level of functioning, rather than on whether an offender resembles the profile of a high frequency violent offender.

Consistent with a dynamic conceptualization of risk, VRS ratings may need to be reassessed after a 2 year period because important situational characteristics may influence risk during the time an offender is in the community. The VRS appears to have a "shelf life" of approximately 2 to 3 years, after which, an offender must be re-assessed. For this reason, traditional comparison studies that do not take into account the ramifications of length of follow up time are not appropriate in the comparison of dynamic and static measures, primarily due to the influence of time on dynamic risk measures.

Over a varied, but longer follow up time, it appears that the VRS provides a substantial advantage over the PCL-R in almost all of the present analyses. Despite the relative stability of the PCL-R over time, it appears that it encounters greater predictive difficulties at total follow up (e.g., this is especially evident in the ROC analysis in which the AUC of the PCL-R approaches chance levels at total follow up time). One possible factor which may provide an obstacle for the PCL-R at longer time periods is its inability to recognize age as an important moderator in violent offending patterns. This is consistent with findings that the PCL-R core personality features (i.e., Factor 1 items)

remain relatively unchanged over the life span, although some criminalized behaviours and propensity for violence may be reduced significantly (Harpur & Hare, 1994).

Conversely, the VRS generally demonstrates a strong improvement in its relationship to violent recidivism at total follow up time. Although the nature of the total follow up time makes this finding difficult to interpret, it is hypothesized that the ability of the VRS to incorporate age may have assisted in improving its predictive abilities as age becomes a more important factor in predicting risk (i.e., as offenders approach middle age).

5.5 Implications and Future Research

The present study did not find that changed VRS scores after the treatment period improved risk predictions. Limitations of the study, including lack of interviewing, inadequate file information and VRS coding which does not incorporate an increase in risk are suggested as potential reasons for not detecting change. However, there is some evidence that post treatment VRS ratings may have provided an advantage in the prediction of presence or absence of violent convictions at 1 year after release, as illustrated by ROC analysis. This may indicate that treatment effects are relatively short lived, perhaps due to a lack of community support and/or a failure by offenders to maintain gains made in treatment.

In addition, the current study also suggests that future risk prediction research using dynamic measures should take into account the change in the power of risk prediction over time. In particular, the present study suggests that a number of low risk offenders may begin to recidivate violently after the two year period. Re-assessment of risk after a two year period may be essential in order to keep the community safe.

Most importantly, the present study suggests that future research comparing static and dynamic measures of risk through comparisons which utilize only a single follow up time are inappropriate. Conclusions based on comparisons which are not sensitive to length of follow up time may be misleading and overly simplistic.

An important focus for future research using the VRS is to determine if re-assessment of dynamic risk items will improve predictions of violent recidivism, that is, to determine if reassessment can improve the predictive abilities of the VRS over longer follow up periods.

Although the present study indicates that the predictive utility of the VRS is greatly impacted by length of follow up time, the nature of this variation deserves future research attention. In particular, the VRS is comprised of both static and dynamic risk sections. The present study has investigated the predictive utility of the VRS by using a VRS total score, which combines the static and dynamic risk items. Future research is encouraged to examine the independent relationship of the static and dynamic sections of the VRS in predicting violent recidivism. In particular, it would be expected that the dynamic risk items of the VRS would experience a peak in predictive ability at 2- 3 years post treatment, and demonstrate a rapid decline thereafter. Conversely, it would be expected that the static items of the VRS would demonstrate a more invariant predictive relationship across follow up time.

In a similar manner, past research has indicated that although the personality factor (i.e., Factor 1) of the PCL-R is relatively stable across the life span, the antisocial behavior factor (i.e., Factor 2) of the PCL-R varies as offenders approach middle age (Harpur & Hare, 1994). Future research exploring the possible dynamic nature of Factor

2 of the PCL-R may also be important in understanding how psychopathy is associated with violence.

The present study suggests that the "overall" nature of the VRS is dynamic, and the "overall" nature of the PCL-R is stable. However, investigation of the contrasting static and dynamic components of these measures may be important to better understand how each relates to violent recidivism.

Finally, the present study suggests that understanding the utility of risk instruments in predicting violent recidivism relies on many features, including the type of recidivism data required, the comprehensiveness of information used in assessments, the time period of concern, and the nature of the risk instrument. In particular, the current study suggests caution in future comparisons of static and dynamic risk measures, and that the length of follow up time after assessment is a vital component which must be incorporated in risk prediction comparisons. Future research utilizing the VRS is strongly urged to incorporate a 2-3 year re-assessment period, in order to investigate the potential advantages and limitations of the VRS. In addition, the use of comprehensive file information and/or the inclusion of interview information also may be necessary to assess accurately the changes in risk level in future research and in practical application of the VRS. It is hoped that this type of future research may provide the necessary guidelines for assisting mental health professionals to make more responsible and accurate risk predictions.

6. References

Andrews, D. A. (1982). The Level of Service Inventory (LSI): The first follow-up. In Toronto: Ministry of Correctional Services.

Andrews, D. A., & Wormith, J. S. (1989). Personality and crime: Knowledge destruction and construction in criminology. Justice Quarterly, 6, 289-309.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. Criminology, 28 (3), 369-465.

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

Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. Psychological Bulletin, 123, 123-142.

Cleckley, H. M. (1976). The mask of insanity (5th ed.). St. Louis, MO: Mosby.

Douglas, K. S., Macfarlane, E. & Webster, C. D. (1996). Predicting dangerousness in the contemporary Canadian mental health and criminal justice systems. Canada's Mental Health, 43(3), 4-11.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works. Criminology, 34, 575-607.

Greenberg, D. F. (1996). Criminal careers. Aldershot, UK: Dartmouth Publishing Company.

Greenwood, P., Rydell, C. P., Abrahamse, A. F., Caulkins, J. P., Chiesa, J., Model, K. E., & Klein, S. P. (1996). Implementing the law. In D. Schichor & D. Sechrest (Eds.), Three strikes and you're out. Thousand Oaks, CA: Sage Publications.

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. Personality and Individual Differences, 1, 111-119.

Hare, R. D. (1981). Psychopathy and violence. In J. R. Hayes, T. K. Roberts, & K. S. Soloway (Eds.), Violence and the violent individual (pp. 53-74). Jamaica, NY: Spectrum.

Hare, R. D. (1991). The Hare Psychopathy Checklist-Revised. Toronto: Multi-Health Systems.

Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. Criminal Justice and Behavior, 23, 25-54.

Hare, R. D., & Jutai, J. W. (1983). Criminal history of the male psychopath: Some preliminary data. In K. T. Van Dusen & S. A. Mednick (Eds.), Prospective studies of crime and delinquency. Boston: Kluwer-Nijhoff.

Hare, R. D., & McPhearson, L. M. (1984). Violent and aggressive behavior by criminal psychopaths. International Journal of Law and Psychiatry, 7, 329-337.

Harpur, T. J., & Hare, R. D. (1994). Assessment of psychopathy as a function of age. Journal of Abnormal Psychology, 103 (4), 604-609.

Harris, G. T., Rice, M. E., & Cormier, C. A. (1989). Violent recidivism among psychopaths and nonpsychopaths treated in a therapeutic community. Journal of Consulting and Clinical Psychology, 57, 374-379.

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.

Hart, S. D., Kropp, P. R., & Hare, R. D. (1988). Performance of psychopaths following conditional release from prison. Journal of Consulting and Clinical Psychology, 56, 227-232.

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. Legal and Criminological Psychology, 3, 139-170.

Kozol, H. L., Boucher, R. J., & Garafalo, R. F. (1970). The diagnosis and treatment of dangerousness. Crime and Delinquency, 18, 371-392.

Martinson, R. (1974). What works? Questions and answers about prison reform. The Public Interest, 35, 22-54.

Martinson, R. (1979). New findings, new views: A note of caution regarding prison reform. Hofstra Law Review, 7, 243-258.

Meehl, P. E. (1954). Clinical versus statistical prediction. Minneapolis, MN: University of Minnesota Press.

Monahan, J. (1980). The clinical prediction of violent behavior. Maryland: National Institute of Mental Health.

Monahan, J. (1996). Violence prediction: The past twenty years and the next twenty years. Criminal Justice and Behavior, 23(1), 107-120.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. Journal of Consulting and Clinical Psychology, 62(4), 783-792.

Nuffield, J. (1982). Parole decision-making in Canada. Ottawa: Communication Division, Solicitor General Canada.

Otto, R. K. (1992). Prediction of dangerous behavior: A review and analysis of "second-generation" research. Forensic Reports, 5, 103-133.

Rice, M. E., & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. Journal of Consulting and Clinical Psychology, 63(5), 737-748.

Serin, R. C. (1991). Psychopathy and violence in criminals. Journal of Interpersonal Violence, 6, 423-431.

Simon, J. (1996). Criminology and the recidivist. In D. Schichor & D. Sechrest (Eds.), Three strikes and you're out. Thousand Oaks, CA: Sage Publications.

Toni, G., Wong, S., & Burt, G. (1999). Assessing treatment changes in violent psychopathic offenders using the Violence Risk Scale. Manuscript in preparation.

Webster, C. D., & Eaves, D. (1993). The HCR-20: A scheme for the assessment of dangerousness and risk. Mental Health, Law and Policy Institute Department of Psychology, Simon Fraser University.

Wong, S. (1984). The criminal and institutional behaviors of psychopaths (No. 1984-87). Ontario, Ottawa, Canada: Ministry of the Solicitor General of Canada.

Wong, S., & Gordon, A. (1996). Violence Risk Scale - Experimental Version 1 (VRS). Department of Psychology and Research. Regional Psychiatric Centre, Saskatoon, SK., Canada: Solicitor General of Canada.

Wong, S., Templeman, R., Gu, D., Andre, G. & Leis, T. (1999). Criminal Career Profile: A quantitative index of past offending. Submitted for publication.

Appendix A

THE PSYCHOPATHY CHECKLIST

Robert D. Hare

Name: _____

FPS NO. _____

Ratings
(0,1,2)

1. Glibness/superficial charm

|_____|_____| _____

2. Grandiose sense of self worth

|_____|_____| _____

3. Need for stimulation/proneness to boredom

|_____|_____| _____

4. Pathological lying

|_____|_____| _____

5. Conning/manipulative

|_____|_____| _____

6. Lack of remorse or guilt

|_____|_____| _____

7. Shallow affect

|_____|_____| _____

8. Callous/lack of empathy

|_____|_____| _____

9. Parasitic lifestyle

|_____|_____| _____

10. Poor behavioural controls

|_____|_____| _____

11. Promiscuous sexual behaviour

|_____|_____| _____

12. Early behavioural problems

|_____|_____| _____

13. Lack of realistic, long-term goals

|_____|_____| _____

14. Impulsivity

|_____|_____| _____

15. Irresponsibility

|_____|_____| _____

16. Failure to accept responsibility for own actions

|_____|_____| _____

17. Many short-term marital relationships

|_____|_____| _____

18. Juvenile delinquency

|_____|_____| _____

19. Revocation of conditional release

|_____|_____| _____

20. Criminal versatility

|_____|_____| _____

Total: _____

	Yes	No
Did you use file information for the above assessment?	[ ]	[ ]
Did you interview the patient for the above assessment?	[ ]	[ ]

Appendix B

VRS Score Sheet

If it is necessary to omit rating a Static or Dynamic Factor, the rater should indicate whether the omission is because there is insufficient information (I) or because the item is not applicable (N).

Static Factors

						I or N
S1	Current Age	0	1	2	3	_____
S2	Age at First Violent Conviction	0	1	2	3	_____
S3	Number of Juvenile Convictions	0	1	2	3	_____
S4	Violence throughout Lifespan	0	1	2	3	_____
S5	Prior Release Failures/Escapes	0	1	2	3	_____
S6	Stability of Family Upbringing	0	1	2	3	_____

Total Static Factor Score Before Treatment: _____

Total Static Factor Score After Treatment: _____

(only if there are changes to S1 or S5)

Dynamic Factors and Total Scores

		Part A	Part B	Total (A+B)	IorN
D1	Violent Lifestyle	0 1 2 3	-1.5 -1 -.5 0	____	____
D2	Criminal Personality	0 1 2 3	-1.5 -1 -.5 0	____	____
D3	Criminal Attitudes	0 1 2 3	-1.5 -1 -.5 0	____	____
D4	Work Ethic	0 1 2 3	-1.5 -1 -.5 0	____	____
D5	Criminal Peers	0 1 2 3	-1.5 -1 -.5 0	____	____
D6	Interpersonal Aggression	0 1 2 3	-1.5 -1 -.5 0	____	____
D7	Emotional Disinhibition	0 1 2 3	-1.5 -1 -.5 0	____	____
D8	Violence During Incarceration	0 1 2 3	-1.5 -1 -.5 0	____	____
D9	Weapon Use	0 1 2 3	-1.5 -1 -.5 0	____	____
D10	Insight into the Cause of Violence	0 1 2 3	-1.5 -1 -.5 0	____	____
D11	Mental Illness	0 1 2 3	-1.5 -1 -.5 0	____	____
D12	Substance Abuse	0 1 2 3	-1.5 -1 -.5 0	____	____
D13	Stability of Relationships With Significant Others	0 1 2 3	-1.5 -1 -.5 0	____	____
D14	Community Support	0 1 2 3	-1.5 -1 -.5 0	____	____
D15	Released to High Risk Situations	0 1 2 3	-1.5 -1 -.5 0	____	____
D16	Violence Cycle	0 1 2 3	-1.5 -1 -.5 0	____	____
D17	Impulsivity	0 1 2 3	-1.5 -1 -.5 0	____	____
D18	Cognitive Distortion	0 1 2 3	-1.5 -1 -.5 0	____	____
D19	Compliance with Community Supervision	0 1 2 3	-1.5 -1 -.5 0	____	____
D20	Security Level of Release Institution	0 1 2 3	-1.5 -1 -.5 0	____	____

Total Dynamic Factor Score			
Total Static Factor Score From Previous Page			
Total Static and Dynamic Factor Score			

Clinical Override _____ _____

Yes **No**

Appendix C
ROC Curve Analysis

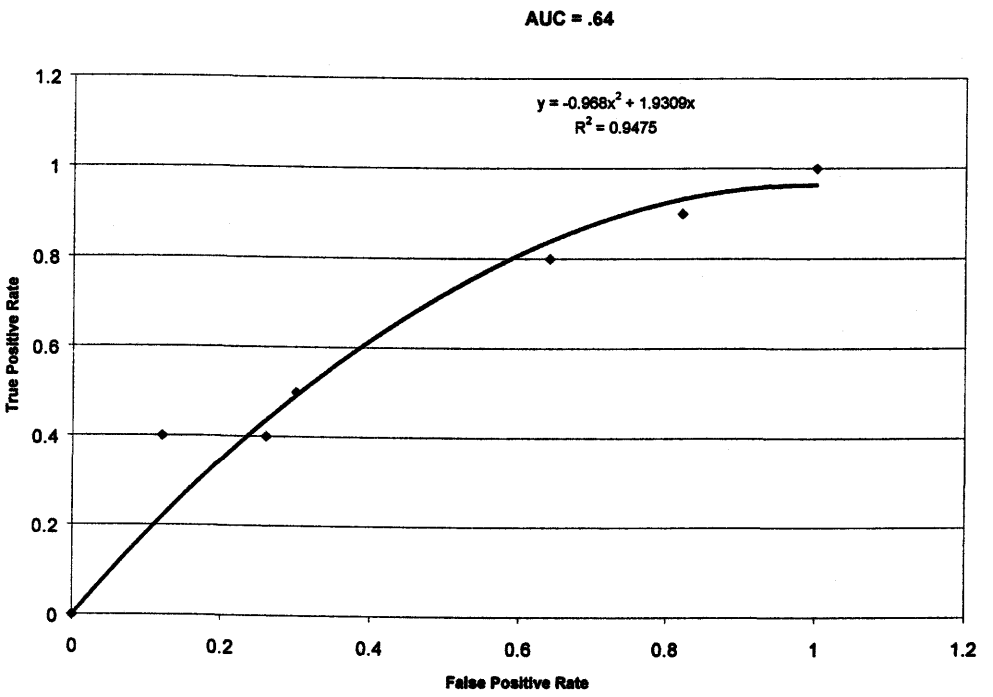

Figure 1 ROC Analysis of PCL-R at 1 Year Post-Release

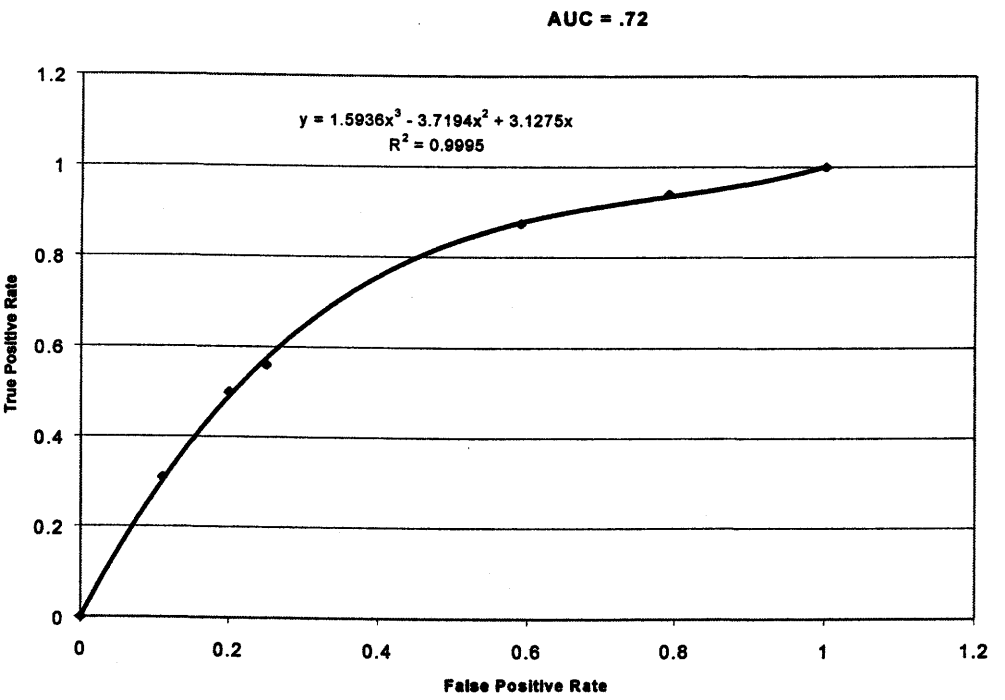

Figure 2 ROC Analysis of PCL-R at 2 Years Post-Release

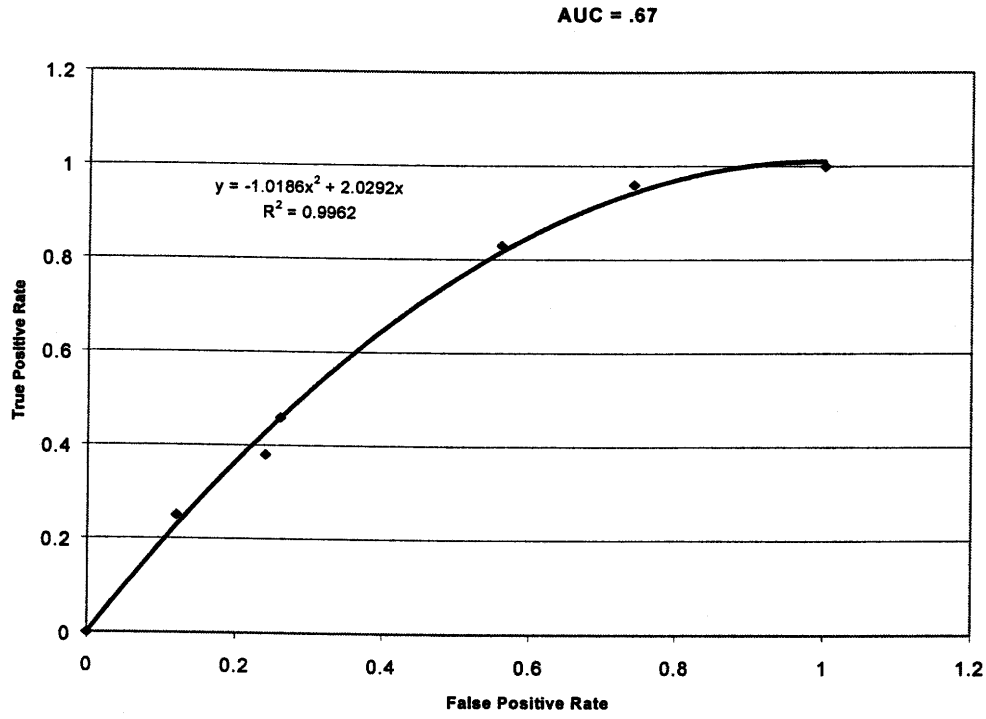

Figure 3 ROC Analysis of PCL-R at 3 Years Post-Release

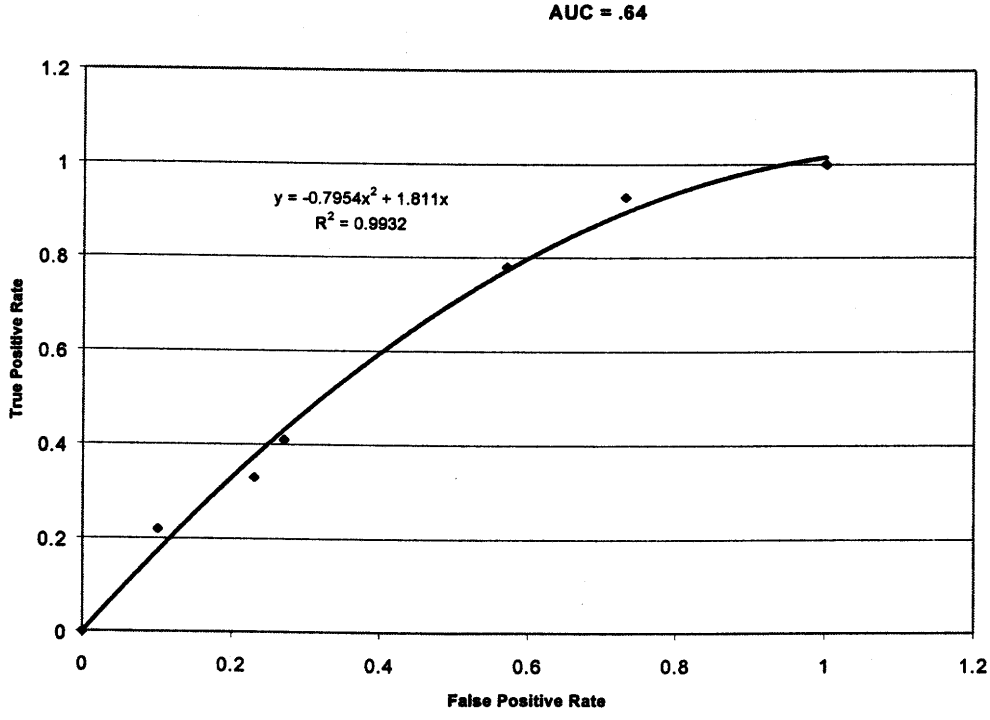

Figure 4 ROC Analysis of PCL-R at 4 Years Post-Release

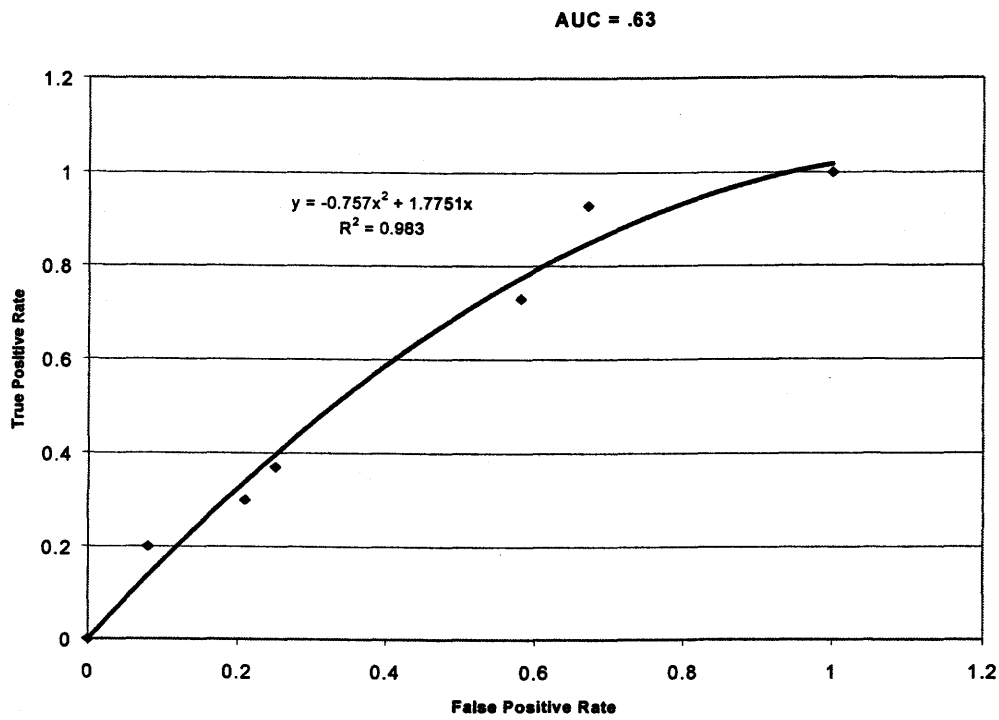

Figure 5 ROC Analysis of PCL-R at 5 Years Post-Release

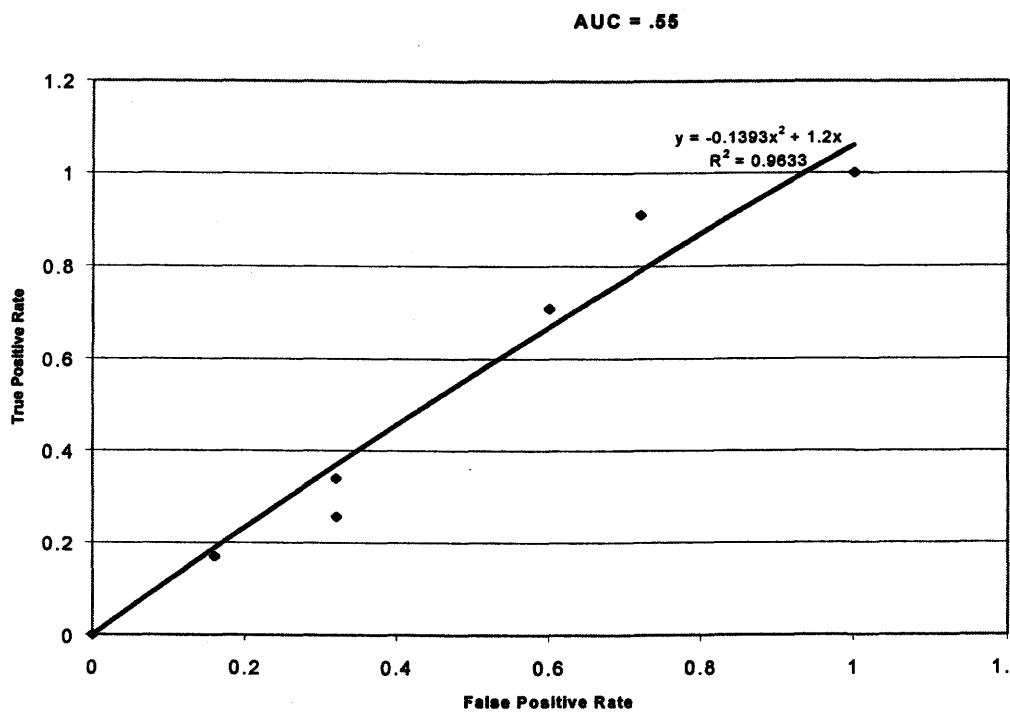

Figure 6 ROC Analysis of PCL-R at Total Follow-Up Time

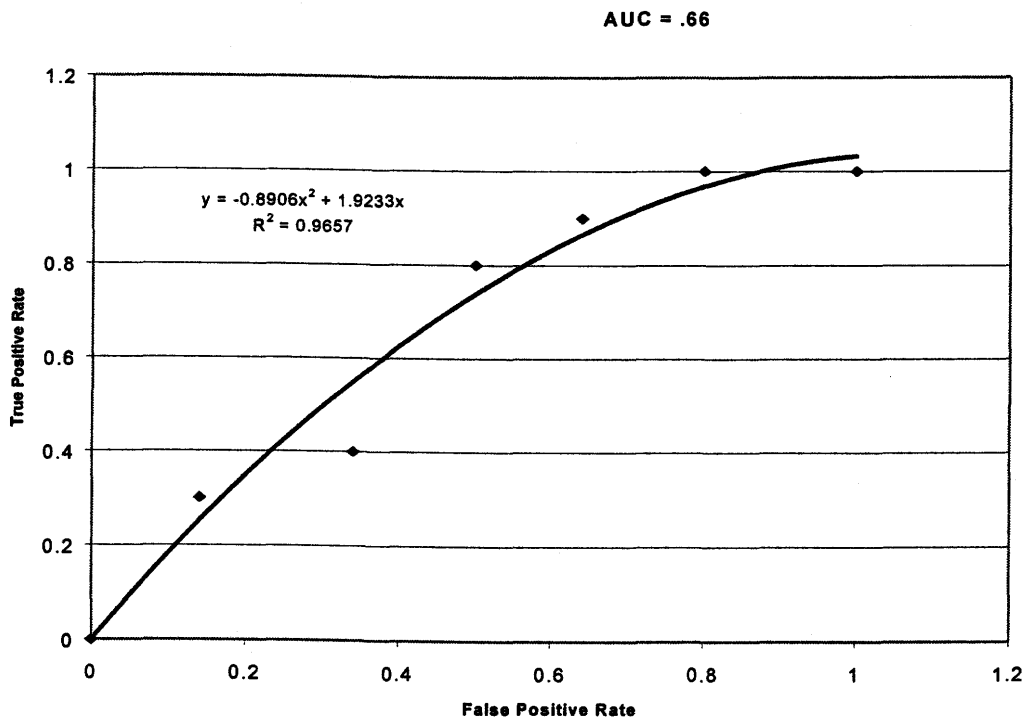

Figure 7 ROC Analysis of VRS Pre-Treatment Ratings at 1 Year Post-Release

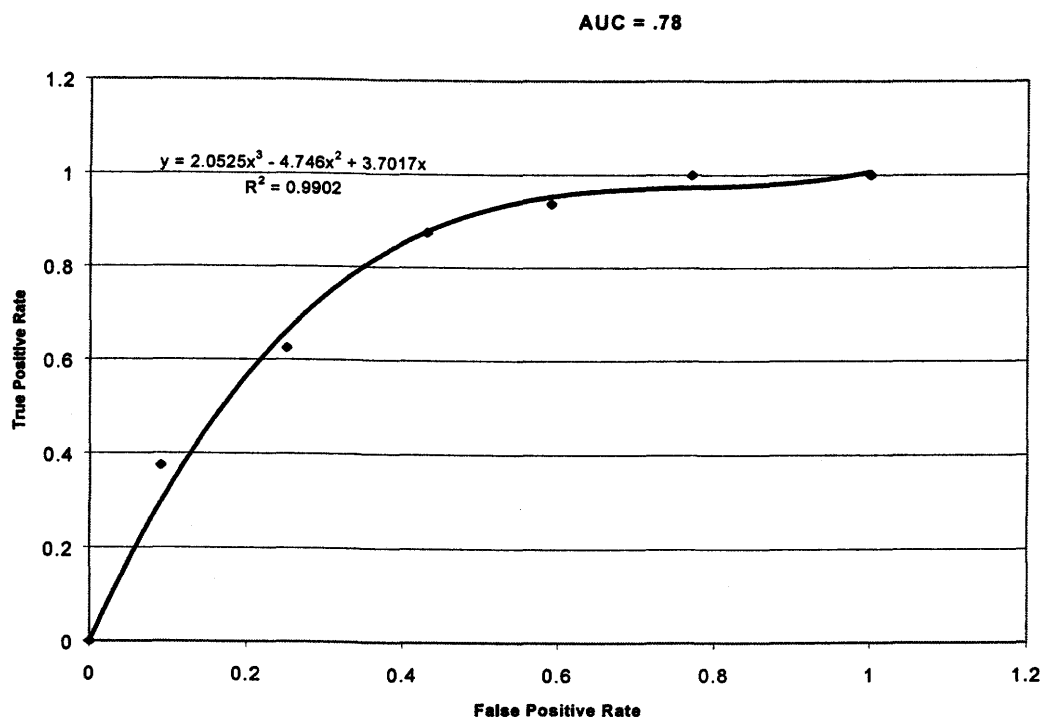

Figure 8 ROC Analysis of VRS Pre-Treatment at 2 Years Post-Release

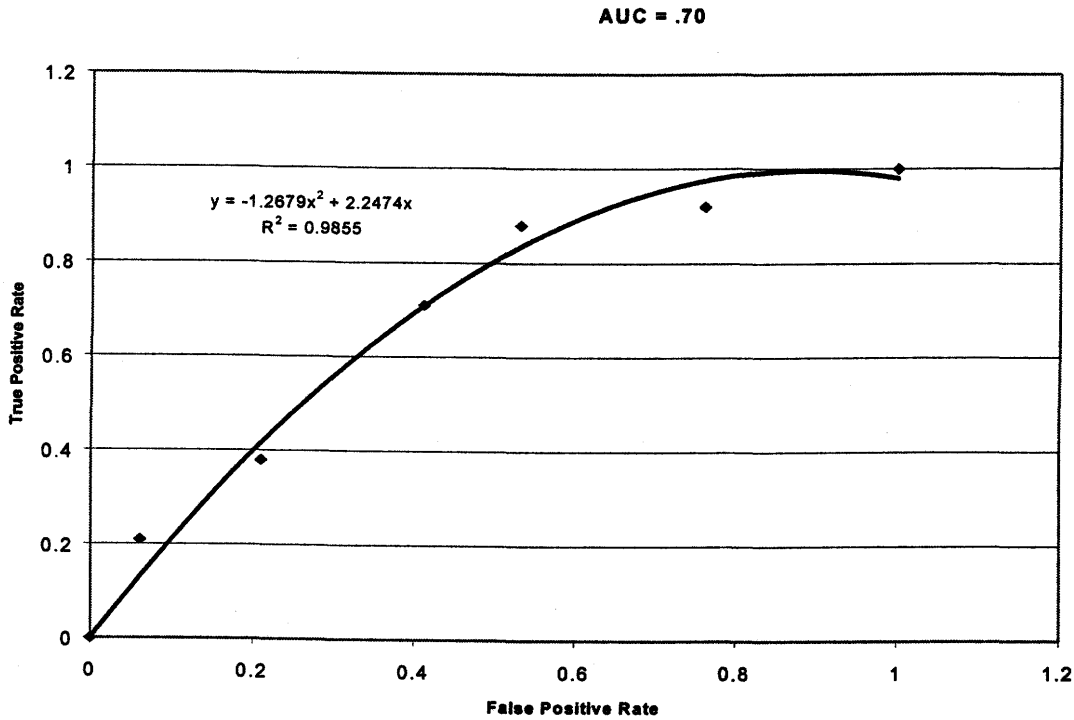

Figure 9 ROC Analysis of VRS Pre-Treatment Rating at 3 Years Post-Release

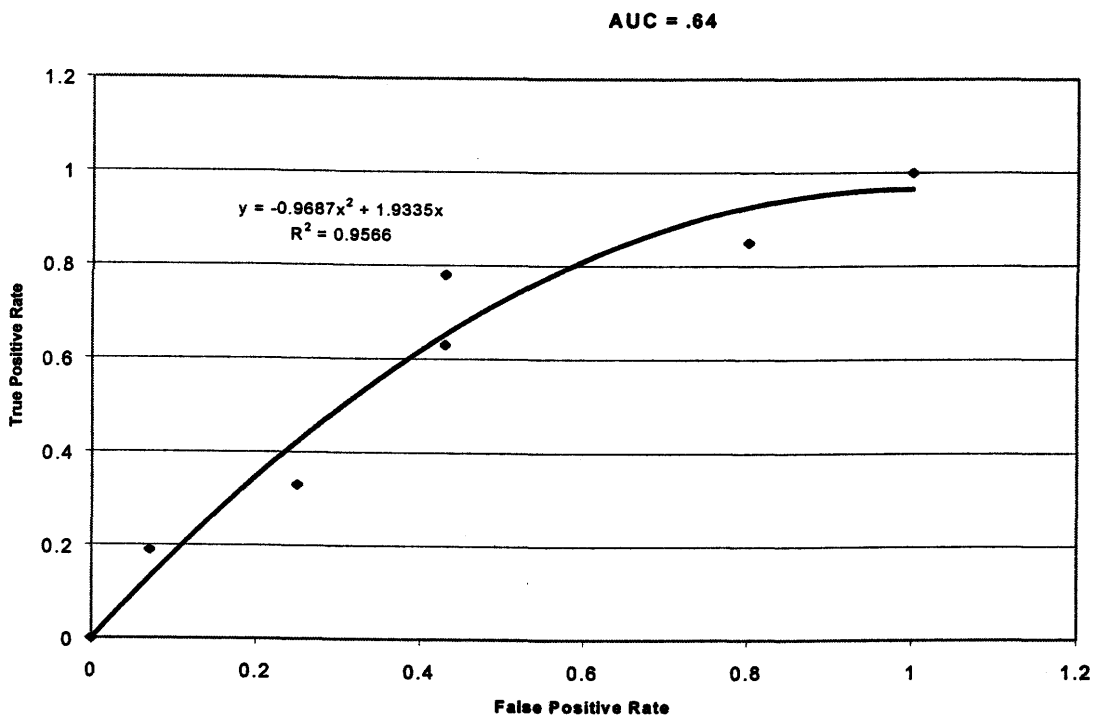

Figure 10 ROC Analysis of VRS Pre-Treatment Rating at 4 Years Post-Release

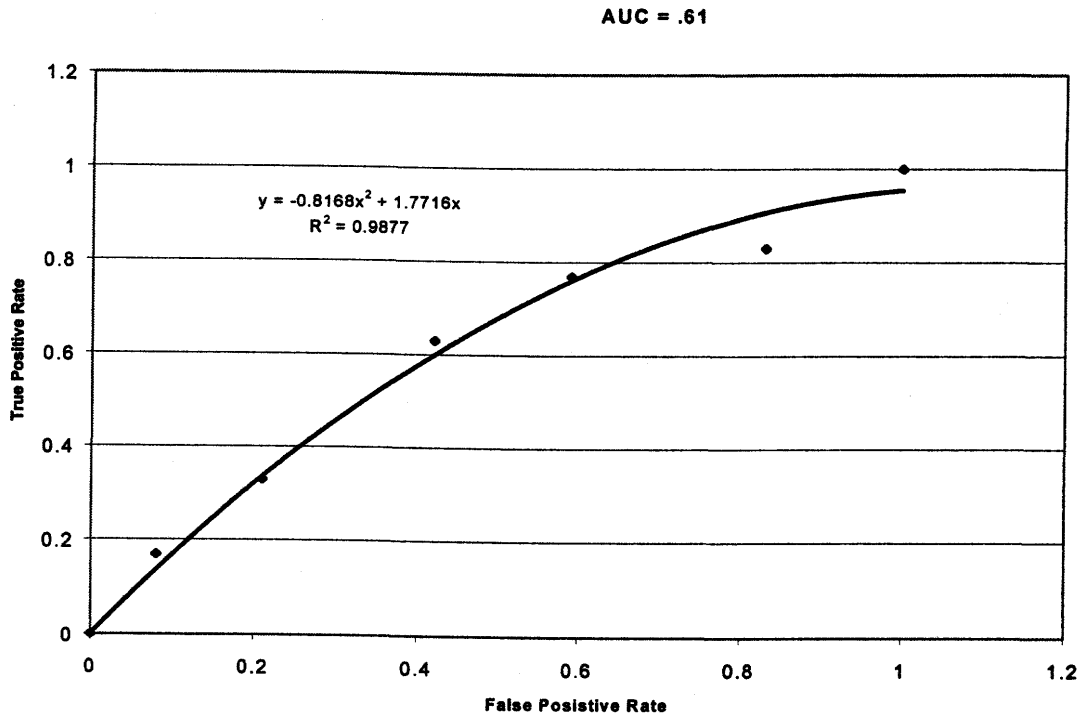

Figure 11 ROC Analysis of VRS Pre-Treatment Rating at 5 Years Post-Release

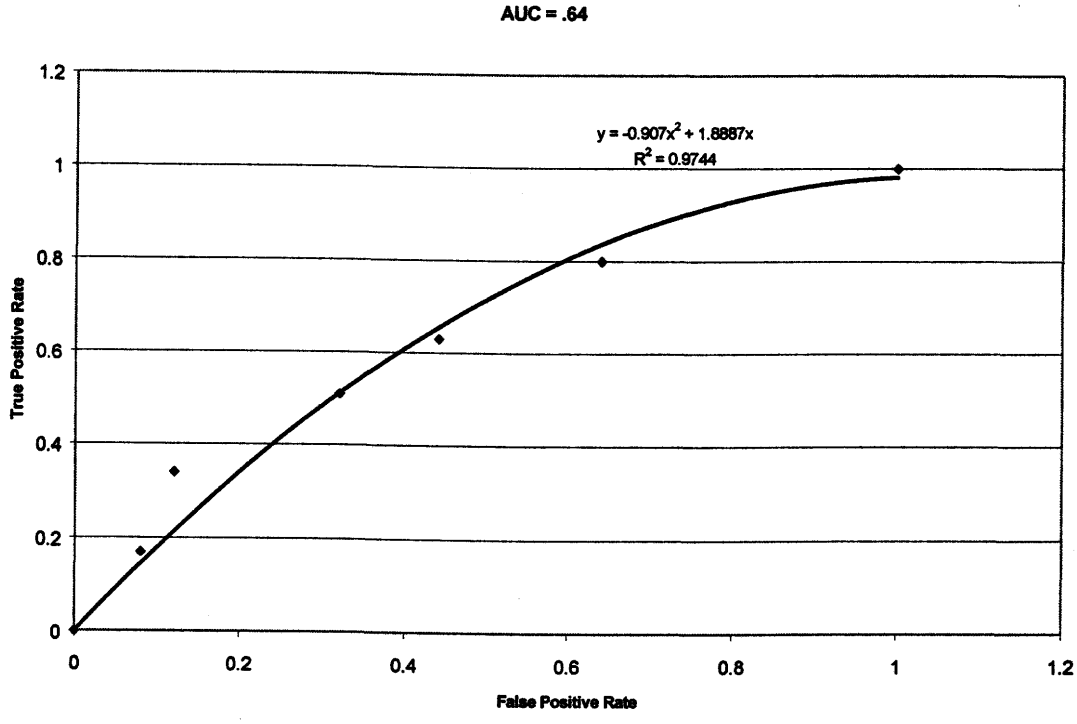

Figure 12 ROC Analysis of VRS Pre-Treatment Rating at Total Follow-Up Time

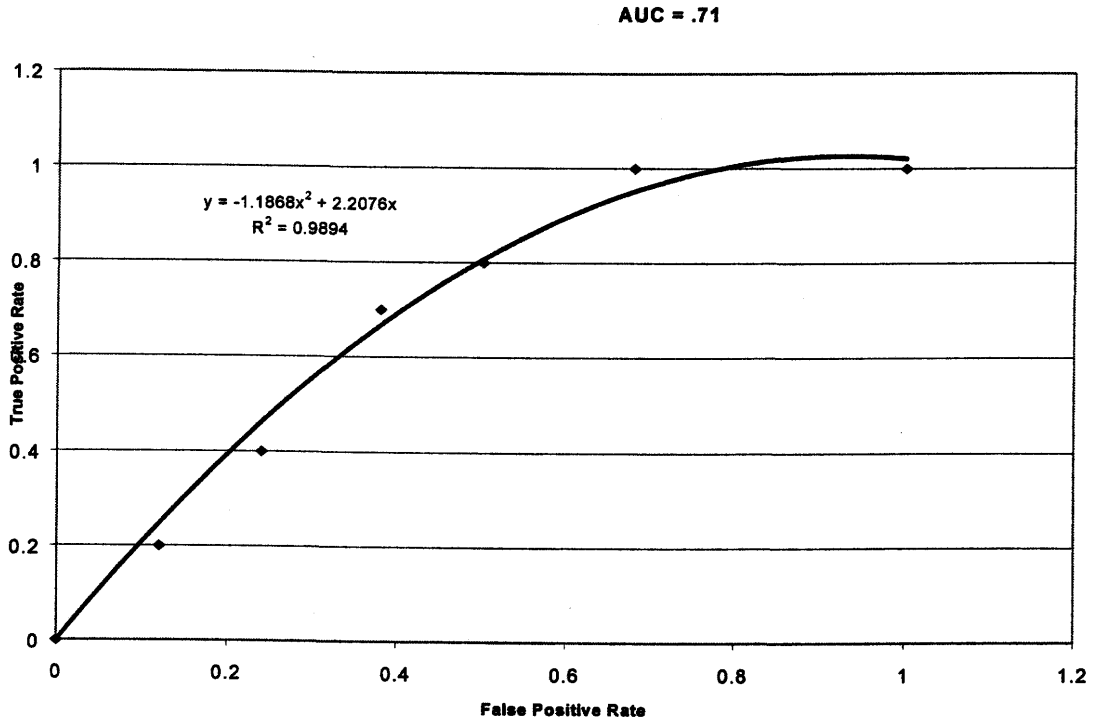

Figure 13 ROC Analysis of VRS Post-Treatment Rating at 1 Year Post-Release

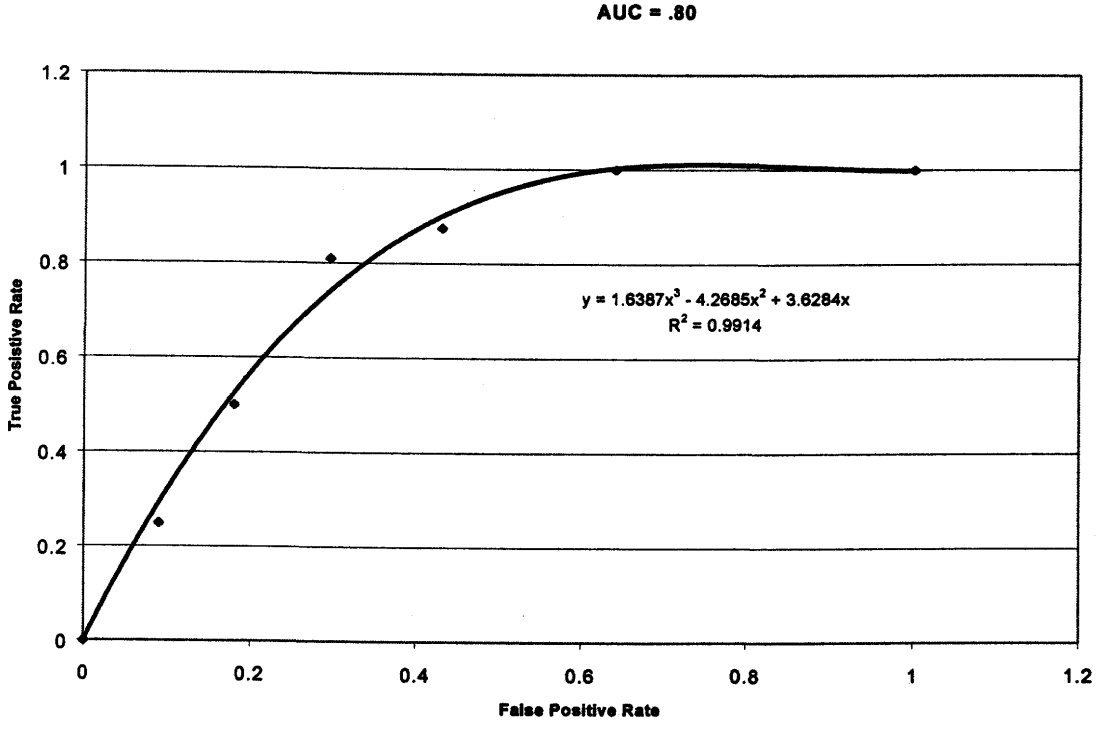

Figure 14 ROC Analysis of VRS Post-Treatment Rating at 2 Years Post-Release

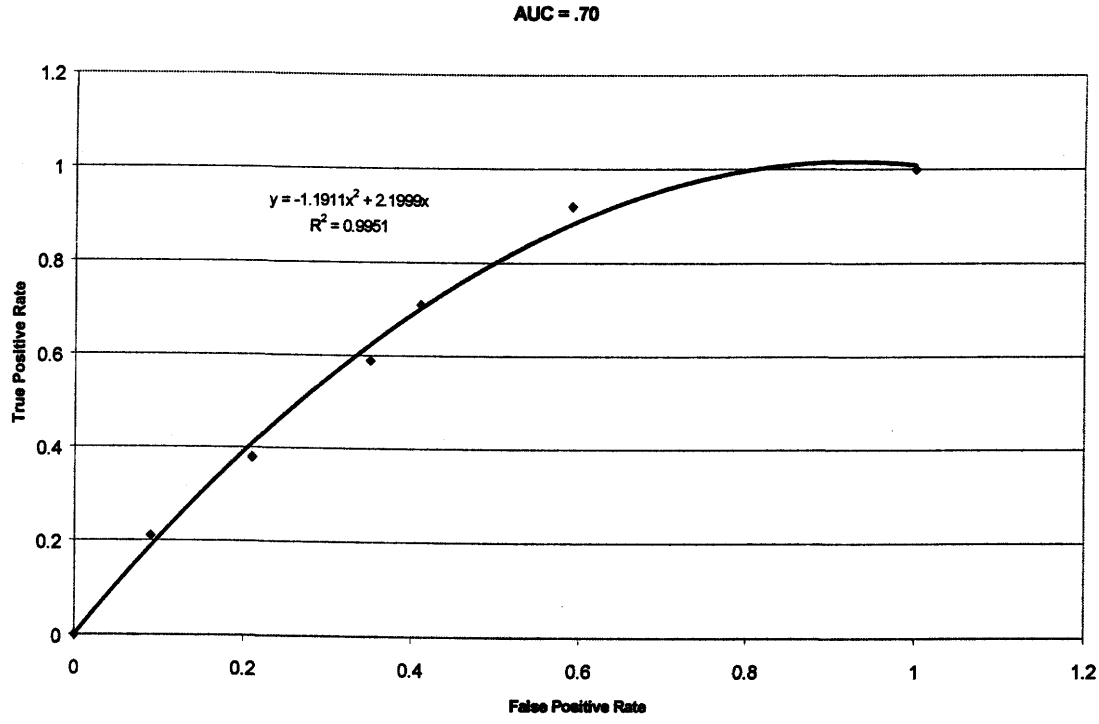

Figure 15 ROC Analysis of VRS Post-Treatment Rating at 3 Years Post-Release

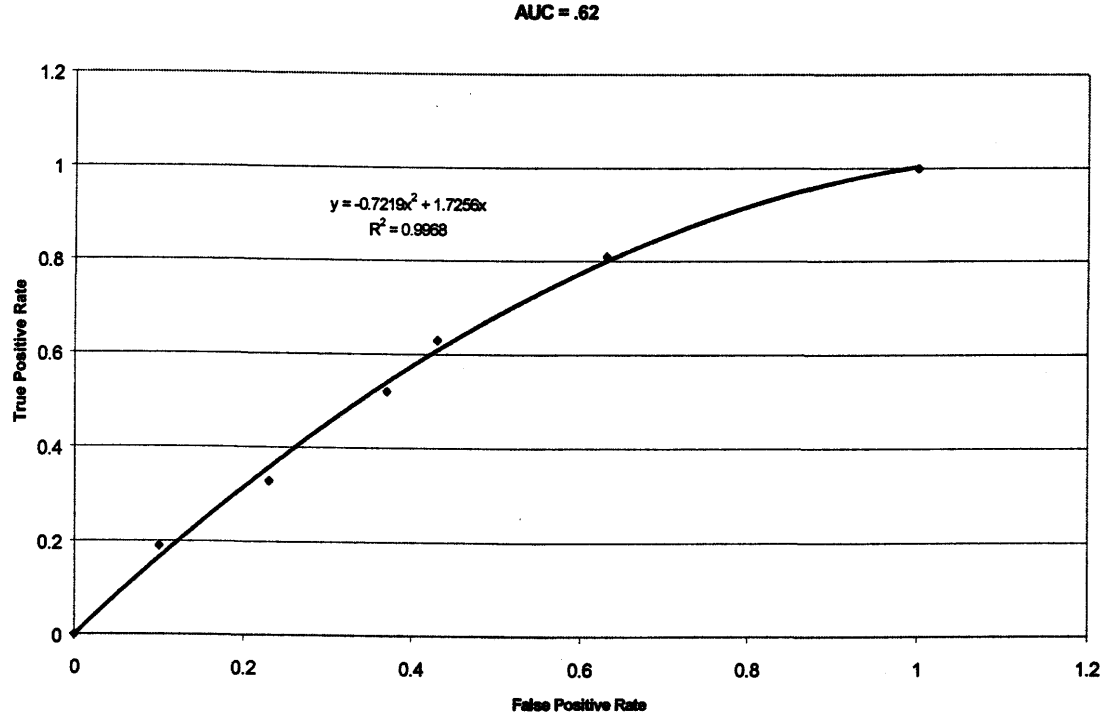

Figure 16 ROC Analysis of VRS Post-Treatment Rating at 4 Years Post-Release

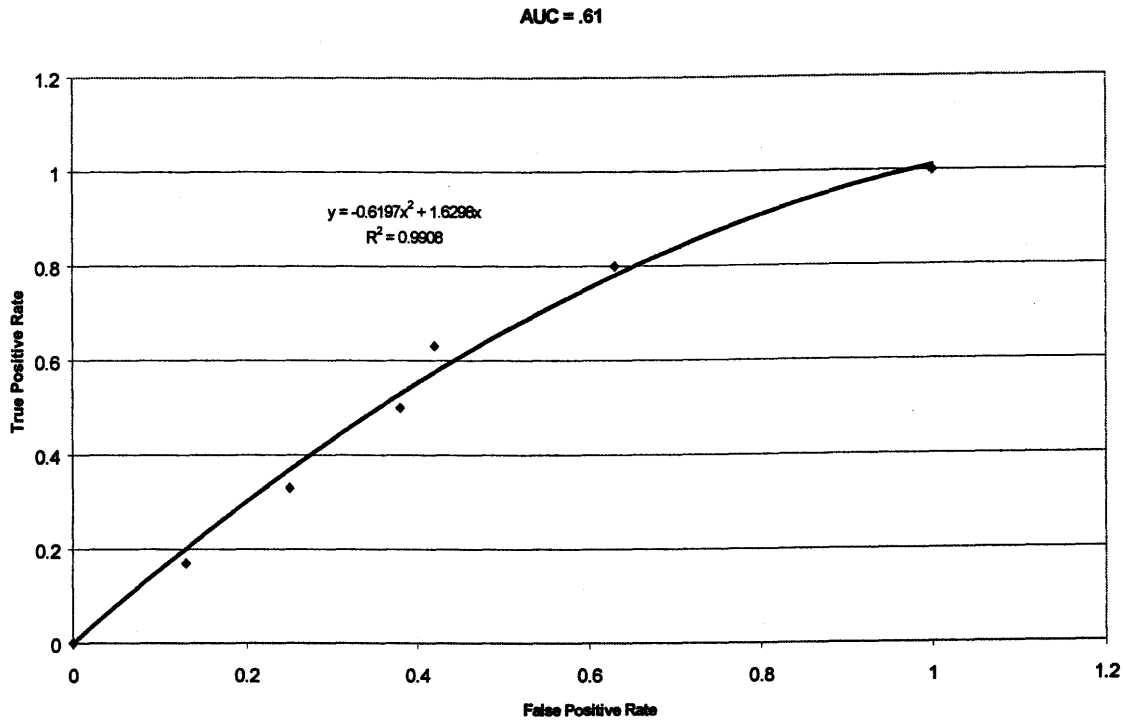

Figure 17 ROC Analysis of VRS Post-Treatment Rating at 5 Years Post-Release

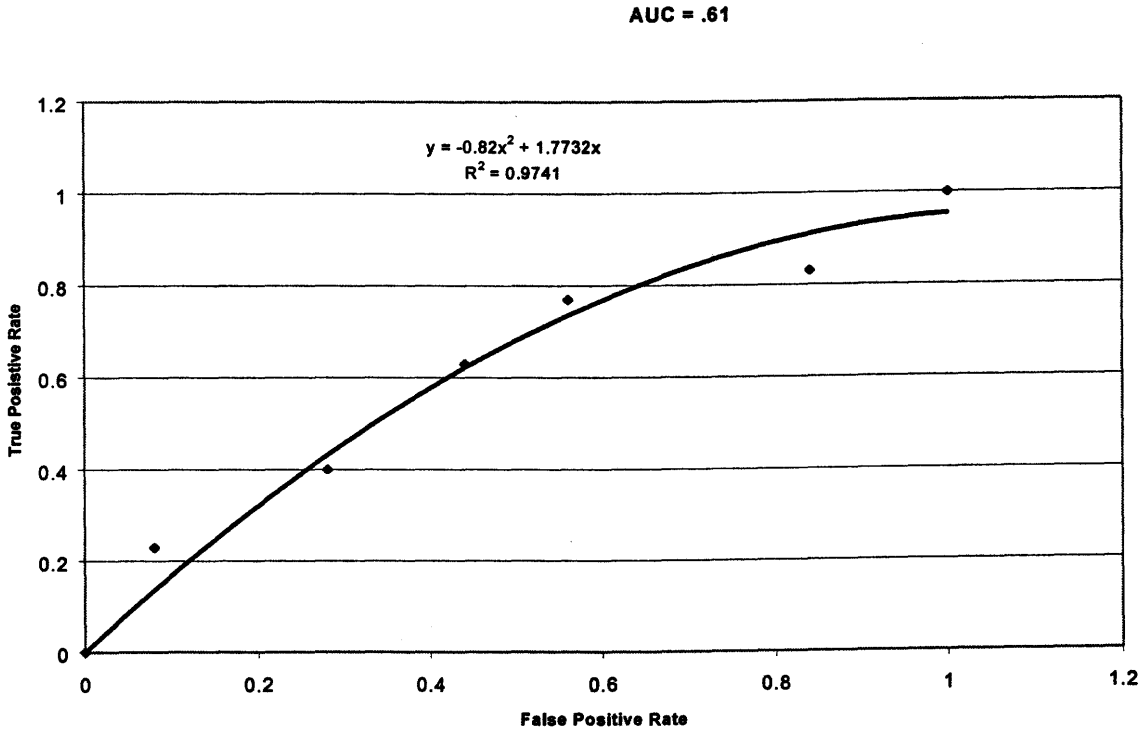

Figure 18 ROC Analysis of VRS Post-Treatment Rating at Total Follow-Up Time

Appendix D

Table 1 Cumulative Percentage of Offenders who Violently Recidivated at 1- 5 years follow up.

Risk Group	Cumulative Follow Up Time				
	Year 1	Year 2	Year 3	Year 4	Year 5
High (PCL-R > 24)	35%	55%	55%	58%	65%
Moderate (PCL-R > 16) (PCL-R < 25)	20%	32%	48%	53%	58%
Low (PCL-R < 17)	10%	16%	21%	32%	44%