THE ASSOCIATION BETWEEN PRESENTENCE RISK EVALUATIONS AND SENTENCING OUTCOME

Sandy Jung and Kelsey Brown
MacEwan University

Liam Ennis
Integrated Threat and Risk Assessment Centre

Denise Ledi
Alberta Hospital Edmonton

Presentence risk evaluations are routinely submitted to judges prior to sentencing with the goal of guiding treatment recommendations and informing sentencing decisions. Empirical research has yet to examine this association. In the present study, the correspondence between presentence risk evaluations and sentencing outcomes was explored using the files of 165 offenders who had completed LSI-R, LS/CMI, and/or HCR-20 protocols. We found that sentencing outcome was associated with risk assessment scores, particularly from the Level of Service Measures. This study's findings suggest that presentence risk evaluations have a strong association with the sentencing decisions of the judiciary, suggesting that sentencing procedures may reflect evidence-based practice.

Keywords: risk assessment, sentencing, incarceration, Level of Service Measures, HCR-20

The inclusion of risk assessments in presentence evaluations is becoming an increasingly common practice in criminal justice and correctional settings. Despite their growing use in bail and pretrial release decisions, in probation decisions, and in predicting future behavior of parolees (Kleiman, Ostrom, & Cheesman, 2007), relatively little is known about the extent to which risk assessments are guiding the sentencing decisions of the judiciary. In this paper, we investigate the association between presentence risk evaluations and sentencing outcomes through a retrospective review of presentence evaluations conducted by trained mental health professionals. We begin with an examination of the issues raised regarding the use of risk assessment in the sentencing process and a review of the limited empirical studies that have examined the judiciary's views on using risk assessments in sentencing hearings.

Correspondence concerning this article should be addressed to Sandy Jung, Department of Psychology, MacEwan University, Edmonton, Alberta, Canada T5J 2P2. E-mail: sandy.jung@macewan.ca

Published work in the fields of criminology and forensic psychology reveals a consensus that the sentencing of criminals has several goals, including deterrence, incapacitation, retribution, rehabilitation, and/or restitution (Greene & Heilbrun, 2011). Most U.S. jurisdictions and Commonwealth countries purport sentencing to be proportionate to the nature of the offense and the level of responsibility assumed by the offender (Criminal Code of Canada, 2013). Therefore, a judge must consider the severity of the index offense and any factors associated with the offender's degree of responsibility when determining a suitable punishment for the crime committed. Furthermore, the sentences assigned should be similar among comparable offenders who have committed related crimes. These principles are employed to decrease disparity among judges; however, this disparity is considered the most significant problem regarding the sentencing process in Canada. According to Silver and Chow-Martin (2002), these disparities are due, at least in part, to judges' reliance on informal and intuitive approaches to making predictions regarding the likelihood of future dangerous or illegal behavior by criminal defendants.

Psychological risk assessments and psychiatric reports submitted to judges prior to sentencing are intended to reduce variability in sentencing decisions, to protect the public, and to guide the allocation of limited correctional and rehabilitative resources. These assessment protocols determine an offender's probability of recidivating on the basis of various legal (e.g., past convictions) and extralegal (e.g., offender age) factors (Pozzulo, Bennell, & Forth, 2012). An abundance of research has consistently found that structured, empirically-based risk assessments demonstrate superior predictive accuracy relative to intuitive or unstructured approaches (Hanson & Morton-Bourgon, 2009; Swets, Dawes, & Monahan, 2000) The perceived practical value of risk assessments in sentencing depends on one's perspective on the purposes of punishment (Kleiman et al., 2007). For example, if the goal of sentencing is for *retribution* or *deterrence*, then the fundamental components of a presentence risk assessment (i.e., opinions regarding likelihood of re-offense and identification of criminogenic needs) are largely irrelevant because the purpose of sentencing is to mete out punishment to the offender in proportion to the amount of harm inflicted on the victims of the crime (or what has been called "just desserts;" Cole, 2007; Singer, 1979). On the other hand, if the goal is *utilitarian* in nature, and the goal of judicial sanctions is the protection of the public through preventative incapacitation and/or rehabilitation, then the issues addressed through structured and valid risk assessment procedures would seem to be highly relevant to judicial decision makers. Structured risk assessments offer a superior alternative to current judicial practices, which base predictions of future risk on subjective impressions and intuitions (Tonry, 1987) and, consequently, risk assessments have the potential to make sentencing more uniform, consistent, and objective (Silver & Chow-Martin, 2002).

Researchers suggest that the future of violence risk assessment should emphasize psychometrically sound considerations of meaningful and relevant risk factors using objective data (Hanson, 2009). Past research has demonstrated that risk evaluations that include empirically validated factors have been associated with increased accuracy in predicting recidivism when compared to methods that rely on professional judgment (Harris, Rice, & Cormier, 2002; Hanson & Morton-Bourgon, 2009). Kleiman et al. (2007) studied a

process developed in Virginia that uses risk assessment measures to redirect nonviolent, prison-bound offenders into community-based sentence programs. Results of their study indicate that, when used in the sentencing process, risk assessment measures assist the judiciary in successfully distinguishing between high and low risk offenders. This provides an unbiased, consistent, and more accurate alternative to determining an offender's risk for reoffending than the common reliance on unstructured professional judgment. Marlowe, Festinger, Lee, Dugosh, and Benasutti (2006) found similar results. They hypothesized that risk assessment measures would be useful in matching participants to an appropriate schedule for judicial status hearings. Results showed that offenders performed significantly better in drug court when risk assessment measures were used to inform the frequency of judicial status hearings. More specifically, when higher risk offenders were matched to biweekly drug courts, results demonstrated better outcomes during treatment, as compared to offenders assigned to regular status hearings.

Although these studies provide positive support for the use of risk assessments in judicial decision-making, validated risk tools continue to be misused and misinterpreted (Hanson, 2009) or disregarded entirely (Hilton & Simmons, 2001). To examine this issue, Tata, Burns, Halliday, Hutton, and McNeill (2008) used a qualitative design to investigate the perspectives of both evaluators and the judicial decision-makers. They found that the judges tended to be overly critical and were unlikely to consider the information during the sentencing process. According to Cole (2007), those opposed to the use of risk assessment measures in the sentencing process reason that evaluators are often insufficiently trained, proportionality should not be overridden by risk, and the reliability and validity of the instruments being used were poor. Despite the presence of some opposition, a national survey in Canada noted that 68% of judges thought that risk assessment information should be included in presentence evaluations, with only 21% of judges who expressly opposed the inclusion of risk assessments (Bonta, Bourgon, Jesseman, & Yessine, 2005).

Silver and Chow-Martin (2002) argue that, in order for criminal justice agencies to manage their resources more efficiently, they should be directing resources to higher risk cases, and this can only be done if risk assessments are completed. Similarly, Hayes and Geerken (1997) purport that efficient use of resources would be best represented by shorter sentences for offenders with the lowest probability of future criminal activity. In the U.S., a national working group was formed, and this group produced guidelines that recommended that judges should "have offender assessment information available to inform their decisions handling risk management and reduction" (Casey, Warren, & Elek, 2011, p. 1). Specifically, they cite Andrew and Bonta's (2010) principles for effective correctional rehabilitation where the first principle asserts that higher risk cases should receive higher intensity services and supervision.

Examining the association between presentence evaluations of risk and judicial decisions for sentencing has tremendous relevance to forensic psychological practice. The absence of a relationship between recidivism risk and sentence severity would indicate a failure of courts to adhere to Andrews and Bonta's (2010) principles for effective rehabilitation (also known as Risk-Need-Responsivity principles, RNR) and, consequently, would

be deleterious to rehabilitative outcomes. Hence, it is expected that when a presentence risk assessment is conducted, the results of that risk assessment should be associated with sentencing outcomes. The present study examines whether scores on risk assessment measures presented in a presentence evaluation correspond with subsequent sentencing decisions. Clinical files of offenders who were referred for a presentence evaluation were reviewed, and particular attention was drawn to evaluations that included an HCR-20 and/or a Level of Service Measure. It was expected that higher scores on risk assessment measures and their subscales would be associated with increased sentence severity, which would suggest that judges' decisions are commensurate with the risk assessment component of presentence report information.

METHODS

Participants

The sample was obtained from a forensic outpatient clinic, and the offenders included in the sample were referred for a presentence assessment following a criminal conviction (Jung, Daniels, Friesen, & Ledi, 2012). Psychiatrists and psychologists with training in the use of risk measures and risk evaluations conducted the assessments. Only presentence evaluation reports that contained a Level of Service Measure (i.e., LSI-R or LS/CMI) and/or a HCR-20 protocol were included in the study. One hundred and sixty-five evaluations were identified, and sentencing outcome information was obtained for each of these cases. Of the total sample, 26 cases had a completed LSI-R, 94 had completed a LS/CMI, and 66 had a completed HCR-20. There were 16 cases where both an LS/CMI and an HCR-20 were completed, and 5 cases where an HCR-20 and a LSI-R were completed.

A majority of the evaluations were conducted on male offenders (79.4%; $n = 131$), while female offenders comprised 20.6% ($n = 34$) of the sample. The mean age of offenders at the time of index offense was 30.9 years ($SD = 10.80$, ranging from 18 to 86). The average educational level of the sample was 10.6 years of schooling ($SD = 2.44$, ranging from 2 to 16 years). A large proportion of the sample's ethnicity was unknown ($n = 74$), but of those whose ethnicity were recorded, 46.2% ($n = 42$) of offenders were Caucasian; 31.9% ($n = 29$) were Aboriginal or Metis; 6.6% ($n = 6$) were Black; and 6.6% ($n = 6$) were Asian. At the time of the offense, a little over half of the offenders were single (55.2%; $n = 91$); did not have children (61.2%, $n = 101$); and were unemployed (50.9%, $n = 84$). The index offense cluster of the offenders (i.e., most recent offense(s) that was being dealt with before the courts at time of their assessment) included robbery (14.5%, $n = 24$); property-related offenses (e.g., theft; 12.1%, $n = 20$); assault/threats (37.0%, $n = 61$); substance-related offenses (6.1%, $n = 10$); fraud (4.8%, $n = 8$); breaking and entering (6.1%, $n = 10$); sexual assault (3.6%, $n = 6$); and other offenses, such as weapons-related, driving-related, and obstruction of justice offenses, collectively (15.8%, $n = 26$).

Measures

Level of Service Measures (LSMs)

Level of Service Measures (LSMs) include either the Level of Service Inventory – Revised (LSI-R; Andrews & Bonta, 1995) or the Level of Service/Case Management

Inventory (LS/CMI; Andrews, Bonta, & Wormith, 2004), which is a revised version of the LSI-R. Both measures are theoretically-based risk assessment measures that consist of static and dynamic risk factors designed to categorize offenders according to risk as well as identify potential treatment needs.

The LSI-R has 54 items that are sorted into 10 subscales (i.e., criminal history, family/marital, education/employment, companions, alcohol/drug problem, accommodation, financial, emotional/personal, attitudes/orientation, and antisocial patterns). Each item is scored as either absent or present with total scores ranging from 0 to 54. A higher score on the LSI-R denotes a higher risk to recidivate. The LSI-R has strong overall interrater reliability (*ICC*s = .80–.94) as well as good internal consistency (Cronbach's alpha = .70). Its predictive validity for general recidivism has a mean *AUC* of .71 and, for violent recidivism, it has a mean *AUC* of .64 (Andrews, Bonta, & Wormith, 2006; *AUC*s calculated using the conversion table of Rice & Harris, 2005). These numbers coincide somewhat with an independent review of the LSI-R (*ICC* = .94; general recidivism, *AUC* = .693; violent recidivism, *AUC* = .667; Kroner & Mills, 2001).

The LS/CMI is a 43-item revised version of the LSI-R. Specifically, items and subscales that were not found to strongly relate to recidivism were omitted from the LSI-R (Andrews & Bonta, 2010). The accommodation, financial, and emotional/personal subscales, as well as a number of items from the criminal history, education/employment, and alcohol/drug problem subscales were omitted. An antisocial pattern subscale was also added in the LS/CMI. The LS/CMI items are mostly identical to the LSI-R items, but given the reduction of both the items (from 54 to 43) and subscales (i.e., accommodations, financial, and emotional/personal subscales were omitted, while antisocial pattern was added, hence reducing the subscales from 10 to 8), the total score of the LS/CMI ranges from 0 to 43. Currently, there are no published studies that have provided internal consistency or interrater reliability for the LS/CMI measure. Andrews et al. (2006) reported the predictive validity for general (*AUC* = .739) and violent recidivism (*AUC* = .666) as fairly strong.

In this study, we included the eight subscales from the LS/CMI. Because some of the data include LSI-R protocols, three subscales were modified to exclude certain items and therefore reflect the subscales as measured on the LS/CMI (i.e., removed items: on the criminal history subscale, escape history and official record of assault; on the alcohol/drug problems subscale, other indicators; on the companions subscale, a social isolate). Seven subscales were calculated from the LSI-R or the LS/CMI, but given that only the LS/CMI protocol contained the antisocial pattern subscale, the total score solely was examined using the LS/CMI. Given these minor adaptations, we refer to the risk/need factors as stemming from the LSM.

Historical Clinical Risk Violence Risk Measure (HCR-20)
The HCR-20 (Webster, Douglas, Eaves, & Hart, 1997; Webster, Eaves, Douglas, & Wintrup, 1995) is a structured assessment tool that integrates 20 items to form three scales: historical, clinical, and risk management. The historical scale consists of 10 items (e.g., previous violence, relationship instability, psychopathy). The clinical scale includes five

items (e.g., lack of insight, unresponsive to treatment). The risk management scale includes five items (e.g., plans lack feasibility, exposure to destabilizers). Although the HCR-20 was intended to be a structured professional guide, a quantifiable scoring scheme was used to produce a total score for each offender, which has been used in other empirical endeavors, where a 0 was assigned when the factor was not present for the offender, 1 when there was some suggestive evidence, and 2 when there was clear evidence for the factor's presence. Using this quantified scoring scheme, the total score for the HCR-20 could range from 0 to 40. Published studies consistently report moderate to strong interrater reliability for both the overall HCR-20 total score (e.g., *ICC*s range from .79 to .91; Dahle, 2006; de Vogel & de Ruiter, 2004; Gray, Taylor, & Snowden, 2008) and the scales (e.g., historical scale, *ICC*s range from .82 to .92; clinical scale, *ICC*s range from .64 to .92; RM scale *ICC*s range from .57 to .85; Dahle, 2006; de Vogel & de Ruiter, 2004; Gray et al., 2004). Internal consistency was also good for the total score of the HCR-20 (Cronbach's alphas range from .85 to .95; Belfrage, 1998; Dahle, 2006) and the historical, clinical, and risk management scales (Cronbach's alphas = .96, .89, and .85, respectively; Belfrage, 1998).

Sentencing Outcome

Data from federal criminal records were coded for two variables to assess sentencing outcomes accurately. For the first variable, sentence type, records were coded as custodial, conditional, or probationary. It is important to note that, in Canada, a conditional sentence order is a disposition that allows an offender to serve what is technically a custodial sentence in the community. If the offender does not abide by community conditions, community placement may be rescinded and the offender must serve the remainder of the sentence in jail. In light of this technical distinction, sentencing outcome was defined as actual incarceration (i.e., were physically sent to jail or prison), and incarcerated offenders were compared with a combined group of offenders who were serving their sentence in the community as the result of either a conditional or probationary sentence. The second variable, length of sentence, was measured in terms of the number of months, and lengths were examined for actual incarceration ($M = 30.5$, $SD = 41.86$), combined actual incarceration and conditional sentence ($M = 27.1$, $SD = 36.14$), community sentence (i.e., conditional and/or probationary sentence; $M = 24.4$, $SD = 12.16$), and probation only ($M = 23.6$, $SD = 9.88$).

Procedure

Clinical files of offenders who were referred for an evaluation prior to sentencing were reviewed. The files and evaluation reports were coded on several variables including offending information and items from risk assessment measures. The risk assessment variables were coded directly from the assessor's raw data used in their presentence evaluation. To ensure that the sentencing outcome was coded reliably, 20 of the files were re-coded independently and blind from the original coding to establish interrater reliability. Intraclass correlations (*ICC*s) were calculated and were in the acceptable range (*ICC*s from .94 to 1.00). Sentencing outcomes were obtained and scored following the file review.

RESULTS

This study's primary objective was to examine whether risk assessment measures were associated with the sentencing decisions of the judiciary. To study the relationship, differences were examined by sentencing outcome. To examine the difference between offenders who did and did not receive actual incarceration (i.e., custodial and conditional sentences), an analysis of variance (ANOVA) was conducted for the total score for each measure by each category of sentence, while a multivariate analysis of variance (MANOVA) was performed on all factors for each risk measure. Hence, for the LSM, an ANOVA was conducted on the total score by sentence type, and a MANOVA was performed on the combined eight risk/need factors of the LSM: criminal history, education/employment, family/marital, alcohol/drug problem, leisure/recreation, companions, attitudes/orientation, and antisocial patterns. For the HCR-20, an ANOVA was conducted on the total score by sentence type, and a MANOVA was conducted on the three scales of the HCR-20: historical, clinical, and risk management. Correlational analyses also were performed to determine the relationship between the risk assessment measures and the length of sentence for each type of sentence. To ensure that our findings were not reflecting the age of the offender, we conducted means comparisons and correlations with the outcome variables (i.e., sentence category and length of sentence), and no significant differences (ps < .05) were noted.

Level of Service Measures

The difference between those offenders who received incarceration (M = 22.24, SD = 9.05) and those who did not (M = 16.03, SD = 7.48) on the total LSM score was analyzed, and a significant difference emerged, showing that those who received incarceration received higher overall risk scores than those who did not, $F(1,89)$ = 11.2, p < .001.

A MANOVA did not reveal a significant difference with the combined dependent variables, Wilks' Lambda = 1.7, *ns*. Univariate analyses, at an α of .05, revealed that offenders in the incarcerated group had higher average scores on the criminal history factor than those who did not get an incarceratory sentence (see Table 1 for means, standard deviations, and univariate means comparison statistics for each factor). Significant differences were found for education/employment, family/marital, alcohol/drug problems, companions, and antisocial patterns factor scores, weighing in the direction of higher scores for those who received actual incarceration.

Table 1

Comparison of offenders on the total score and the eight factors of the LSM by incarceration sentence

LSM total and factors	Incarcerated	Not incarcerated	F	η^2
Total score	22.24 (9.05)	16.03 (7.48)	11.2***	.11
Criminal History	4.43 (2.27)	2.88 (2.18)	13.1***	.10
Education/Employment	4.95 (2.85)	3.37 (2.50)	9.2**	.07
Family/Marital	1.86 (1.25)	1.37 (1.07)	4.5*	.04
Alcohol/Drug Problems	4.43 (2.33)	2.86 (2.47)	12.0***	.09
Leisure/Recreation	1.47 (0.70)	1.30 (0.74)	1.6	.01
Companions	2.25 (1.34)	1.37 (1.33)	11.8***	.09
Attitude/Orientation	1.09 (1.00)	0.79 (1.01)	2.5	.02
Antisocial Patterns	1.91 (1.27)	1.27 (1.33)	5.2*	.06

*$p < .05$, **$p < .01$, and ***$p < .001$. Means and standard deviations in parentheses are listed. Incarcerated offenders ranged from an n of 76 to 77, and not incarcerated offenders had an n of 43, except for univariate analysis for antisocial patterns (ns of 58 and 33, respectively). Entering all 8 factors into a multivariate analysis of variance was not significant, Wilks' Lambda = 1.7, ns, $\eta^2 = .14$.

To examine the relationship between each factor of the LSM and the length of sentence given, Pearson's correlation coefficients were calculated (see Table 2 for correlations). It was predicted that the subscale scores and the total score of the LSM would be positively correlated with the length of the sentence in the direction where those who receive higher scores on the LSM factors and on the total LSM score would receive longer incarceration sentences than their lower risk counterparts. Our results indicated that there were no significant associations between each LSM factor and sentence length for actual incarceration, community sentence, or probation only. However, a single correlation emerged between the total LSM score and the sentence length of technical incarceration (i.e., either actual incarceration or conditional sentence).

HCR-20 Total and Scale Scores
When the HCR-20 was examined between incarcerated and non-incarcerated offenders, a difference was found, $F(1,44) = 4.9$, $p < .05$, $\eta^2 = .10$, showing that offenders who received a custodial sentence had higher scores on average ($M = 20.3$, $SD = 6.95$) than offenders who received a community disposition ($M = 15.6$, $SD = 7.12$). No difference was found using a MANOVA on the combined three scales of the HCR-20, Wilks' Lambda = 2.8, $p = .053$, although probability approached significance (see Table 3 for means, standard deviations, and univariate analyses). Univariate analyses revealed a single significant difference, indicating that those who were actually incarcerated scored higher on the historical scale than those who did not receive incarceration.

Table 2

Correlations between the LSM, HCR-20, and length of sentence

Risk measure totals, factors, and scales	Length			
	Incarceration	Technical incarceration	Community sentence	Probation
LSM total	0.25	0.25*	0.17	0.12
Criminal History	0.15	0.18	0.10	-0.02
Education/Employment	0.18	0.19	0.08	0.04
Family/Marital	0.02	0.05	-0.01	0.08
Alcohol/Drug Problems	0.15	0.16	0.05	-0.04
Leisure/Recreation	0.16	0.13	0.05	0.03
Companions	0.19	-0.18	0.13	-0.03
Attitude/Orientation	0.06	0.05	0.16	-0.17
Antisocial Patterns	0.18	0.17	0.09	0.08
HCR-20 total	0.34	0.36*	0.11	0.12
Historical	0.32	0.32	-0.01	0.13
Clinical	0.08	0.12	0.28	0.18
Risk Management	0.36*	0.36*	0.08	0.01

For the LSM factors, *n*'s ranged from 58 to 77 for incarceration length, 79 to 104 for technical incarceration sentence length (includes conditional sentence), 58 to 76 for community sentence (includes conditional and/or probation), and 44 to 57 for probation length. For the HCR-20 scales and total, *n*'s range from 29 to 39 for incarceration length, 34 to 50 for technical incarceration sentence length, 31 to 46 for community sentence length, and 26 to 33 for probation length.

Table 3

Comparison of offenders on the total score and the three scales of the HCR-20 by incarceration sentence

HCR-20 total and scales	Incarcerated ($n = 29$)	Not incarcerated ($n = 17$)	F	η^2
Total score	20.3 (6.95)	15.59 (7.12)	4.9*	.10
Historical	11.2 (3.62)	8.5 (2.83)	7.3**	.14
Clinical	4.5 (2.15)	3.2 (2.63)	3.4	.07
Risk Management	4.6 (2.50)	3.9 (2.93)	0.7	.02

*$p < .05$, **$p < .01$, and ***$p < .001$. Means and standard deviations in parentheses are listed. For incarceration, multivariate analysis of variance was not significant, Wilks' Lambda = 2.8 $p = .053$, $\eta^2 = .17$.

The relationship between each scale of the HCR-20 and the total HCR-20 score with the length of sentence was examined using Pearson's correlation coefficients (Table 2). Results indicated that there were no significant associations between the historical or clinical scales and the sentence length. Significant positive correlations emerged between the risk management scale and the length of incarceration assigned—whether it was actual incarceration, $r(36) = 0.36$, $p < .05$, or technical incarceration (includes actual incarceration or conditional sentence), $r(46) = 0.36$, $p < .05$. Also, a positive correlation emerged between the HCR-20 total score and the sentence length of technical incarceration, $r(32) = 0.36$, $p < .05$.

DISCUSSION

This study's findings suggest that there is a relationship between the results of structured risk assessment procedures and the type of sentence handed down in the legal proceedings. For both of the risk assessment schemes that were investigated (i.e., LSMs and HCR-20), higher overall risk scores were associated with greater likelihood of receiving a custodial sentence. For the LSM factors, incarcerated offenders received higher scores on six of the eight factors than community sentenced offenders, and these include criminal history, education/employment, family/marital, alcohol/drug problems, companions, and antisocial patterns subscales. Of the three HCR-20 scales, only the historical scale was notably higher among those who received actual custodial sentences than those who received community sentences. The results of the present study imply that, when evidence-based risk assessments are used, the risk scores correspond with judges' decisions on whether to incarcerate an offender. The LSM domains, also deemed the 'Central 8' risk factors (Andrews & Bonta, 2010), appear to play an important role in differentiating incarcerated and non-incarcerated offenders but only historical information was relevant when the HCR-20 was used.

Although most LSM factors were associated with the decision to incarcerate an offender, none of these factors were associated with the length of the assigned sentence. Previous research suggests that length of a sentence is based largely on the severity of the index offense, but our results indicate that both total scores on validated risk measures (LSM and HCR-20) and the risk management scale of the HCR-20 demonstrated a positive relationship with sentence length. For the HCR-20, the risk management scale was associated with both the length of actual incarceration and technical incarceration. This result suggests that judges may have been influenced by risk assessment information that reflected barriers to risk management in the community.

Our findings are encouraging and support a correspondence between risk assessment outcome and judicial sentencing decisions. Previous research has not addressed the influence of risk assessment measures on the sentencing decisions of the judiciary, despite the fact that the predictive validity of these actuarial risk tools has been established and it has been recommended these tools should be used at sentencing (Casey et al., 2011). Single factor models have solely been investigated. Consistent with our findings, previous studies have concluded that prior criminal record and current offense severity are associated with

sentencing outcomes (Crow, 2008; Vigorita, 2001). Not surprising, of all the subscales of the LSM, criminal history is the best predictor of recidivism for offenders (Girard & Wormith, 2004). It has been pointed out that, as it relates to sentencing, criminal history is actually multi-dimensional and comprises the number of previous convictions, the similarity between previous offenses and the index offense, the severity of prior offenses, the nature of prior sentences, the frequency of offending, the decay of prior convictions, the offender's age at the time of offense, and the implications of prior record in determining the offender's character (Crow, 2008). In keeping with this complex view of criminal history, Vigorita (2001) identified an interaction between type of index offense (i.e., violent vs. non-violent) and criminal history. More specifically, overall criminal record had a more significant effect for nonviolent offenders; whereas, prior criminal offending only affected sentencing severity to a minimal degree for violent offenders. Related findings from Crow (2008) suggest that previous violent offenses had the most significant influence on the probability of incarceration, as compared to less serious offenses, such as property crime and drug possession. The findings presented by Vigorita (2001) also reveal that the effect of prior criminal record differs between jurisdictions, with risk of incarceration increasing the likelihood of incarceration in urban environments.

Despite the significance of the criminal history, research still emphasizes the importance of assessing several risk factors that are predictive of reoffending. Moreover, although judges may be correct in placing significant weight on an offender's criminal history, it is unclear whether they are extracting this information from empirically validated actuarial risk factors or criminal records. It is unlikely these methods are equivalent. It is notable that extant studies have not examined risk as assessed with a standardized and validated measure of risk of recidivism. Most studies have primarily included archival coding of files or reports rather than risk assessment measures that were completed at the time of the presentence assessment. Hence, little is available on the use of risk factors beyond the static variable of criminal history. Perhaps the difficulty may lie with the infrequency of using validated measures. For example, a study by Roehl and Guertin (2000) examined the use of risk assessment measures in the sentencing of offenders who committed intimate partner violence and found that validated measures were less commonly used, despite the availability of such measures.

Research shows that the value in risk assessment measures resides in planning and delivering effective service to offenders, which may ultimately reduce recidivism (Andrews et al., 2006). In particular, the use of risk assessment measures to guide the sentencing process should strengthen adherence with principles of effective treatment and expedite clinical supervision to ensure public protection from reoffending behaviors (Andrews et al., 2006). Consequently, the linkage between risk assessment and programming will be rewarding in both theory and practice. Monahan and Skeem (2014) have recently highlighted the need to incorporate validated risk assessments in sentencing procedures in light of the limited resources in the current state of corrections. Given that services must often wrestle with budgetary constraints and fiscal demands that result from high rates of incarceration

(Kleiman et al., 2007), using risk assessments can more effectively direct resources to those offenders who need it the most.

Skeem (2013) has called for empirical examination of the use of risk assessment technology in the sentencing practices of the judiciary, and the current study provides a starting point to examine judicial sentencing practices when presentence assessments include formal and standardized risk evaluations. Making it policy to include standardized, risk assessments into sentencing practices and processes is highly beneficial and highlights evidence-based sentencing that is already integrated into the sentencing procedure in at least three states in the U.S., including Virginia, Missouri, and Pennsylvania (Hyatt, Bergstrom, & Chanenson, 2011; Warren, 2010).

Although the results of this study have many important implications with respect to sentencing decisions, certain methodological issues must be considered. One of the major limitations is the archival nature of this study, which does not allow for the control of all variables that may have affected the sentencing decisions of the judiciary. A range of peripheral factors have been shown to influence sentencing outcomes, including geographical and legal jurisdiction, offense-specific sentencing guidelines, whether or not the offender entered a guilty plea, and even the extent to which the court's resources were taxed by heavy caseloads (Ulmer, Eisenstein, & Johnson, 2010). Importantly, our research methods are inferential; we can only infer the influence on the judge's decision-making processes from correlational analyses and means comparisons between offenders who were incarcerated in contrast to those who received community sentences. We cannot be sure about what factors were valued by judges in determining an appropriate sentence. Furthermore, the conclusions and scores from the LS measures may not have been explicitly described in the psychological reports submitted to judges. Unfortunately, the quality of the psychological report is determined by the knowledge, skill, and motivation of the report writer, which cannot be controlled in this study (Tata et al., 2008). Additionally, even when high-quality reports are submitted, the judge determines the relevance of the report and may limit the extent to which he/she embraces the sentencing advice in such evaluations (Tata et al., 2008). Moreover, one cannot be certain that judges have read the entirety of the evaluation.

Examining whether risk assessment measures correspond with sentencing decisions of the judiciary is necessary to survey the practical reality of implementing risk evaluations in the judicial decision-making process. The literature well supports the use of validated risk factors and standardized risk assessments to evaluate and predict reoffending behavior, but a limited number of studies have examined its use in the decisions made throughout the various stages within the criminal justice system. Our study contributes to the existing literature by examining the association between presentence risk evaluations and judicial sentencing decisions. The results suggest that psychological reports that include risk assessments correspond with the sentencing decisions of the judiciary. As the sentencing process can be a contentious area in the criminal justice system, greatly affecting both the offender and the general public, it is important to identify if empirically validated measures, and their factors, are associated with such decisions.

Our research provides support for the association between risk assessment and sentencing, but not all factors may have an impact. Our research calls attention to a continued need to educate the important players within the criminal justice system to ensure that empirically validated measures and variables are recognized and the outcomes are considered in the sentencing process. Making sentencing decisions that are based on valid and relevant information could ensure that sentencing is appropriate, effectively uses limited resources, and better protects the community.

REFERENCES

Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). Cincinnati, OH: Anderson.

Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory – Revised.* Toronto, Canada: Multi-Health Systems.

Andrews, D. A, Bonta, J. L., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency, 52*, 7-27. doi:10.177/0011128705281756

Andrews, D. A., Bonta, J. L., & Wormith, J. S. (2004). *Level of Service/Case Management Inventory (LS/CMI): User's manual.* Toronto, Canada: Multi-Health Systems.

Belfrage, H. (1998). Implementing the HCR-20 scheme for risk assessment in a forensic psychiatric hospital: Integrating research and clinical practice. *Journal of Forensic Psychiatry, 9,* 328-38. doi: 10.1080/09585189808402200.

Bonta, J. L., Bourgon, G., Jesseman, R., & Yessine, A. K. (2005). *Presentence reports in Canada* (User Report 2005-03). Ottawa, Canada: Public Safety and Emergency Preparedness Canada.

Casey, P. M., Warren, R. K., & Elek, J. K. (2011). *Using offender risk and needs assessment information at sentencing: Guidance for courts from a national working group.* Williamsburg, VA: National Center for State Courts.

Cole, D. P. (2007). The umpires strike back: Canadian judicial experience with risk-assessment instruments. *Canadian Journal of Criminology and Criminal Justice, 49,* 519-529. doi: 10.3138/cjccj.49.4.493

Criminal Code of Canada, R.S.C. (2013) c. C-46, s. 718.

Crow, M. S. (2008). The complexities of prior record, race, ethnicity, and policy: Interactive effects in sentencing. *Criminal Justice Review, 33,* 502-523. doi: 10.1177/0734016808320709

Dahle, K. P. (2006). Strengths and limitations of actuarial prediction of criminal reoffence in a German prison sample: A comparative study of LSI-R, HCR-20 and PCL-R. *International Journal of Law and Psychiatry, 29,* 431-442. doi: 10.1016/j.ijlp. 2006.03.001.

de Vogel, V., & de Ruiter, C. (2004). Differences between clinicians and researchers in assessing risk of violence in forensic psychiatric patients. *Journal of Forensic Psychiatry and Psychology, 15,* 145-164. doi: 10.1080/14788940410001655916.

Girard, L., & Wormith, J. S. (2004). The predictive validity of the Level of Service Inventory-Ontario Revision on general and violent recidivism among various offender groups. *Criminal Justice and Behavior, 31,* 150-181. doi: 10.1177/0093854803261335

Gray, N. S., Taylor, J., & Snowden, R. J. (2008). Predicting violent reconvictions using the HCR-20. *British Journal of Psychiatry, 192,* 384-387. doi: 10.1192/bjp.bp.107.044065.

Greene, E., & Heilbrun, K. (2011). *Wrightman's psychology and the legal system* (7th ed.)*.* Belmont, CA: Wadsworth.

Hanson, R. K. (2009). The psychological assessment of risk for crime and violence. *Canadian Psychology, 50,* 172-182. doi: 10.1037/a0015726

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21,* 1-21. doi: 10.1037/a0014421

Harris, G. T., Rice, M. E., & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior, 26,* 377-394. doi: 10.1023/A:1016347320889

Hayes, H. D., & Geerken, M. R. (1997). The idea of selective release. *Justice Quarterly*, *14*, 353-370. doi: 10.1080/07418829700093361

Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behavior, 25,* 393-408. doi: 10.1023/A:1010607719239

Hyatt, J. M., Bergstrom, M. H., & Chanenson, S. L. (2011). Follow the evidence: Integrate risk assessment into sentencing. *Federal Sentencing Reporter, 23,* 266-268. doi: 10.1525/fsr.2011.23.4.266

Jung, S., Daniels, M. K., Friesen, M., & Ledi, D. (2012). An examination of convergent constructs among Level of Service Measures and other measures. *The Journal of Forensic Psychiatry & Psychology*, *23,* 601-619. doi: 10.1080/14789949.2012.732595

Kleiman, M., Ostrom, B. J., & Cheesman, F. L. (2007). Using risk assessment to inform sentencing decisions for nonviolent offenders in Virginia. *Crime & Delinquency, 53,* 106-131. doi: 10.1177/0011128706294442

Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior, 28,* 471-489. doi: 10.1177/009385480102800405

Marlowe, D. B., Festinger, D. S., Lee, P. A., Dugosh, K. L., & Benasutti, K. M. (2006). Matching judicial supervision to clients' risk status in drug court. *Crime and Delinquency, 52,* 52-76. doi: 10.1177/0011128705281746

Monahan, J., & Skeem, J. (2014). Risk redux: The resurgence of risk assessment in criminal sanctioning. *Federal Sentencing Reporter, 26,* 158-166. doi: 10.1525/fsr.2014.26.3.158

Pozzulo, J., Bennell, C., & Forth, A. (2012). *Forensic psychology* (3rd ed.). Toronto, Canada: Pearson.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies ROC area, Cohen's *d* and *r*. *Law and Human Behavior, 29,* 615-620. doi: 10.1007/510979-005-6832-7

Roehl, J., & Guertin, K. (2000). Intimate partner violence: The current use of risk assessments in sentencing offenders. *The Justice System Journal, 21,* 171-198. doi: 10.1080/23277556.2000.10871281

Silver, E., & Chow-Martin, L. (2002). A multiple models approach to assessing recidivism risk: Implications for judicial decision making. *Criminal Justice and Behavior*, *29*, 538-568. doi:10.1177/009385402236732

Singer, R. (1979). *Just deserts: Sentencing based on equality and desert.* Cambridge, MA: Ballinger.

Skeem, J. (2013). Risk technology in sentencing: Testing the promises and perils [Commentary on "Actuarial sentencing: An 'unsettled' proposition" by Hannah-Moffat, 2011]. *Justice Quarterly, 30,* 297-303. doi: 10.1080/07418825.2012.687513

Swets, J., Dawes, R., & Monahan, J. (2000). Psychological science can improve diagnostic decision. *Psychological Science in the Public Interest, 1*, 1-26. doi: 10.1111/1529-1006.001

Tata, C., Burns, N., Halliday, S., Hutton, N., & McNeill, F. (2008). Assisting and advising the sentencing decision process: The pursuit of 'quality' in pre-sentence reports. *The British Journal of Criminology, 48,* 835-855. doi: 10.1093/bjc/azn055

Tonry, M. (1987). Prediction and classification: Legal and ethical issues. In D. M. Gottfredson & M. Tonry (Eds.), *Prediction and classification: Criminal justice decision making* (pp. 367-413). Chicago, IL: University of Chicago Press.

Ulmer, J. T., Eisenstein, J., & Johnson, B. D. (2010). Trial penalties in federal sentencing: Extra-guidelines factors and district variation. *Justice Quarterly, 27,* 561-592. doi: 10.1080/07418820902998063

Vigorita, M. S. (2001). Prior offense type and the probability of incarceration: The importance of current offense type and sentencing jurisdiction. *Journal of Contemporary Criminal Justice, 17,* 167-193. doi: 10.1177/1043986201017002006

Warren, R. K. (2010). Evidence-based sentencing: Are we up to the task? *Federal Sentencing Reporter, 23,* 153-158. doi: 10.1525/fsr.2010.23.2.153

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *The HCR-20: Assessing the risk for violence* (Version 2). Burnaby, Canada: Mental Health, Law and Policy Institute, Simon Fraser University.

Webster, C. D., Eaves, D., Douglas, K. S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk* (Version 1). Burnaby, Canada: Mental Health, Law and Policy Institute, Simon Fraser University.